



Europäisches Patentamt
European Patent Office
Office européen des brevets



Veröffentlichungsnummer: **0 571 880 A1**

(12)

EUROPÄISCHE PATENTANMELDUNG

(21) Anmeldenummer: 93108104.6

(51) Int. Cl.⁵: **C12N 9/00, C12Q 1/68,
C12N 15/52**

(22) Anmeldetag: 18.05.93

(30) Priorität: 23.05.92 DE 4217134

(43) Veröffentlichungstag der Anmeldung:
01.12.93 Patentblatt 93/48

(84) Benannte Vertragsstaaten:
**AT BE CH DE DK ES FR GB GR IE IT LI LU NL
PT SE**

(71) Anmelder: **BOEHRINGER MANNHEIM GMBH**

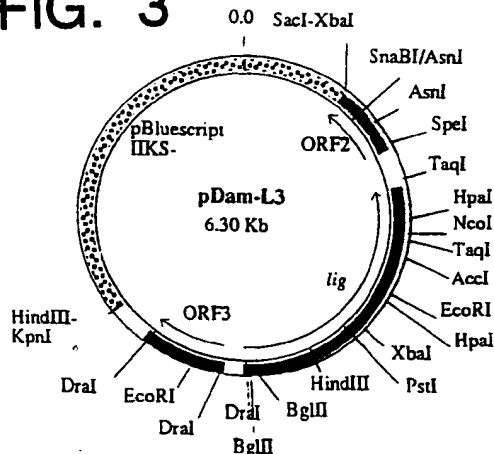
D-68298 Mannheim(DE)

(72) Erfinder: **Kletzin, Arnulf, Dr.**
Agnes-Bernauer-Strasse 172
W-8000 München 21(DE)
Erfinder: **Rüger, Rüdiger, Dr.**
Tutzing Strasse 2
W-8124 Seeshaupt(DE)
Erfinder: **Kessler, Christoph, Dr.**
Schlossbergweg 11
W-8021 Dorfen(DE)
Erfinder: **Kaletta, Cortina, Dr.**
Seehauserstrasse 14
W-8000 München 70(DE)
Erfinder: **Jarsch, Michael, Dr.**
Unterkarpfensee 11
W-8173 Bad Heilbrunn(DE)

(54) **Thermostabile Ligase aus Archaeobakterien.**

(57) Thermostabile DNA-Ligase erhältlich aus Archaeobakterien und Verfahren zu ihrer Verwendung in der Nukleinsäurediagnostik

FIG. 3



Gegenstand der Erfindung ist eine thermostabile Ligase aus Archaeobakterien, Nukleinsäuren, welche für diese Ligase kodieren, und ein Verfahren zum Nachweis von Nukleinsäuren unter Verwendung dieser Ligase.

Ganz allgemein sind Ligasen Enzyme, welche bereits bestehende Enden von Polynukleotiden verknüpfen (zusammenfassend: P. Higgins, N.R. Cozzarelli 1989 in *Recombinant DNA Methology* S. 3-24). Bei einer Ligasereaktion wird die Bildung einer Phosphodiesterbindung zwischen einer freien 5'-Phosphatgruppe und einer im gleichen Strang eines doppelsträngigen DNA-Moleküls gegenüberliegenden, freien 3'-Hydroxylgruppe katalysiert. Die Ausbildung der Phosphodiesterbindung ist gekoppelt an die Spaltung einer Pyrophosphatbindung in einem für die Reaktion notwendigen Co-Faktor. Dabei sind sowohl intra- als auch intermolekulare Reaktionen möglich. Ligasen spielen eine wichtige Rolle bei der DNA-Replikation und bei der Reparatur von DNA-Schäden. Man unterscheidet prinzipiell zwischen zwei Typen von DNA Ligasen. Der erste Typ ist aus Eukaryonten und Bacteriophagen der T-Serie bekannt und benötigt ATP als Cofaktor. Im Gegensatz hierzu sind die Ligasen aus Eubakterien (zweiter Typ) NAD-abhängig. Bekannte Beispiele für bakterielle Ligasen sind zum einen die E-coli-Ligase, und für Phagenligasen die Ligase des Bakteriophagen T4.

Die Ligasereaktion läuft prinzipiell in drei Schritten ab:

- Adenylierung des freien Enzyms (dabei wird bei Co-Faktor NAD ein Nikotinamidmononukleotid und bei Co-Faktor ATP ein Pyrophosphat freigesetzt). Für die genannten Beispiele liegt nach dieser Reaktion AMP über die Epsilon-Gruppe eines Lysin-Restes gebunden vor.
- Der AMP-Rest wird im zweiten Schritt vom Enzym zur 5'-Phosphatgruppe des freien DNA-Endes transferiert, wobei eine energiereiche Pyrophosphatbindung gebildet wird.
- Im dritten Schritt der Ligasereaktion wird über einen nukleophilen Angriff der gegenüberliegenden 3'-Hydroxylgruppe eine Phosphatdiesterbindung ausgebildet und die AMP-Gruppe entfernt.

Eine thermostabile Ligase ist aus *Thermus thermophilus* (J. Biochem. 100, 123-132 (1986)) bekannt. Das Enzym hat ein Molekulargewicht von ca. 79.000 (J. Biol. Chem. Vol. 259, 16, S. 10041 bis 10047 (1984)). Dieses Enzym benötigt Magnesium 2^{+} oder Mangan 2^{+} sowie NAD als Cofaktoren, jedoch kein ATP. Das Enzym aus *Thermus thermophilus* HB 8 wurde auch erfolgreich in der Oligonukleotidligation der Ligase Chain Reaction (LCR) eingesetzt (WO 91/17239). Eine thermostabile, ATP-abhängige DNA-Ligase war bisher nicht bekannt.

Aufgabe der vorliegenden Erfindung war es, eine thermostabile, ATP-abhängige Ligase zu finden. Bevorzugt sollte diese die Nachteile des Standes der Technik nicht aufweisen. ATP-Abhängigkeit ist nämlich gegenüber NAD-Abhängigkeit wegen der Thermostabilität von ATP bevorzugt.

Gegenstand der Erfindung ist daher eine thermostabile Ligase erhältlich aus Archaeobakterien. Ebenfalls Gegenstand der Erfindung sind Nukleinsäuresequenzen, welche für diese Ligase kodieren, ein Verfahren zum Nachweis von Nukleinsäuren mit Hilfe dieser Ligase und ein Verfahren zur Isolierung eines Enzyms.

Eine Ligase im Sinne der Erfindung ist ein Enzym, welches zwei benachbart an einen gemeinsamen Strang anhybridisierte DNA-Einzelstränge miteinander über eine kovalente Phosphodiesterverbindung verknüpft. Ligasen schließen somit Einzelstrangbrüche in Nukleinsäuredoppelsträngen unter Verknüpfung einer 5'-Phosphatgruppe einer Nukleinsäure mit der 3'-Hydroxylgruppe einer anderen Nukleinsäure.

Die erfindungsgemäße DNA-Ligase wird in Archaeobakterien codiert, insbesondere in der Ordnung der Sulfobacterales und im speziellen in *Desulfurolobus ambivalens* (Bergey's Manual, Vol. 3, S. 2250-2253 (1984), Nature Vol. 313, S. 789-791 (1985), System. Appl. Mikrobiol. 8, 197-203 (1986)). *Desulfurolobus ambivalens* gehört genau wie eine größere Anzahl weiterer extrem thermophiler, ungefähr zur gleichen Zeit isolierter Organismen zu dem erst kurz zuvor neu definierten Urreich der Archaeobakterien (FEMS Microbiol. Rev. 75, 117-124 (1990)). Ende der siebziger Jahre hatten sich die Befunde gehäuft, daß die "Prokaryonten" nach molekularen Kriterien in zwei fundamental verschiedene Organismengruppen zerfallen. Dies veranlaßte 1977 C.R. Woese und G.E. Fox, die Prokaryonten als phylogenetische Einheit aufzutrennen, indem sie einen Teil davon als gleichberechtigtes Urreich der Archaeobakterien den übrigen, den Eubakterien, gegenüberstellten und so die bis dahin als phylogenetisch bedeutsam angesehene Prokaryonten-Eukaryonten-Dichotomie durch eine Trichotomie ersetzen. Seither ist eine große Anzahl von bis dahin unbekannten Archaeobakterien beschrieben worden, und auch die Erkenntnisse auf den Gebieten der Molekularbiologie und Biochemie der Archaeobakterien haben das Bild der Gruppe als monophyletische Einheit des Lebens gefestigt (J. Mol. Evol. 24, 167-173 (1986)). Mit der 1990 vorgeschlagenen Neubenennung der Domänen der Eucarya, Archaea und Bacteria als höchste Taxa anstelle der Urreiche der Eukaryonten, Archaeobakterien und Eubakterien wurde dieser Weiterentwicklung Rechnung getragen.

Desulfurolobus ambivalens aus der Ordnung der Sulfobacterales wächst bei einer Temperatur von 55-85 °C und in einem Bereich von pH 0,8-4 (Optima 80 °C und pH 2,5). Es ist strikt chemolithoautotroph und kann auf einem reinen Mineralmedium unter CO₂-Assimilation durch Reduktion oder Oxidation von Schwefel

wachsen.

Desulfurolobus ambivalens ist aus der öffentlichen Sammlung der DSM unter Nr. 3772 erhältlich. Aufgrund der nachfolgend gegebenen Identifikation des Enzyms kann das native Enzym nach bekannten Methoden (z.B. Proteinisolierung, Trennung nach Molekulargewicht oder Ladung und Aktivitätstest) aus *Desulfurolobus ambivalens* isoliert werden.

Die erfindungsgemäße Ligase ist bevorzugt zwischen 560 und 750 Aminosäuren lang. Besonders bevorzugt beträgt die Länge des erfindungsgemäßen Enzyms ca. 600 Aminosäuren.

Die erfindungsgemäße Ligase ist abhängig von ATP als Cofaktor. Die Aminosäuresequenz der DNA-Ligase aus *Desulfurolobus ambivalens* ist in FIG 1 angegeben. Die Homologie dieser Aminosäuresequenz mit eukaryonten DNA-Ligasen beträgt nur ca. 30-35 % (FIG 4). Die Homologie zu Phagenligasen ist noch niedriger. Nennenswerte Homologien zu NAD-abhängigen Ligasen aus Eubakterien konnten (mit Ausnahme des AMP-bindenden Motivs) nicht gefunden werden. Die erfindungsgemäße DNA-Ligase enthält daher entweder die Aminosäuresequenz der FIG 1 oder einen Teil dieser Aminosäuresequenz, der mehr als 50 Aminosäuren lang ist und Ligasenaktivität aufweist, oder eine Aminosäuresequenz, die zu mindestens 40 %, bevorzugt mindestens 60, besonders bevorzugt mindestens 90 % homolog zu einer der genannten Sequenzen ist und Ligaseaktivität aufweist. Sie hat ferner bevorzugt weniger als 30 % Homologie zu der Aminosäuresequenz von eukaryonten DNA-Ligasen. Das Enzym hat bevorzugt ein Molekulargewicht zwischen 60 und 80 kD, besonders bevorzugt von ca. 67 kD. Der isoelektrische Punkt des Proteins aus *D. ambivalens* (aus der Aminosäuresequenz berechnet) ist 6.21. Die Gesamtladung beträgt -6. Der Gehalt an hydrophoben Aminosäuren (I + L + V + M) ist 26.2 %, was innerhalb des Bereiches anderer bekannter DNA-Ligasegene liegt (20.6 - 26.8 %).

Das erfindungsgemäße Enzym ist bei Temperaturen von über 55 °C, bevorzugt zwischen 75 und 95 °C, besonders bevorzugt bei etwa 80 °C, stabil. Insbesondere ist das Enzym stabil unter Bedingungen, welche die mehrmalige Behandlung von Nukleinsäuren unter thermisch denaturierenden Bedingungen beinhalten. Besonders bevorzugt ist das Enzym säurestabil von pH 4 bis pH 8.5, bevorzugt im Bereich von pH 6-7.5. Das erfindungsgemäße Enzym ist aktiv mindestens in dem Intervall von 50-85 °C, insbesondere zwischen 50 und 60 °C.

Größere Mengen des Enzyms können auch durch rekombinante Herstellung gewonnen werden. Dazu wird eine für das Enzym kodierende Nukleinsäure in einen Vektor eingebracht und in einem geeigneten Wirt zur Expression gebracht. Die erforderlichen Arbeitsschritte für die molekulare Klonierung eines Gens sind prinzipiell aus Sambrook, Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press (1989) bekannt. Die für die Ligase kodierenden Nukleinsäuren (DNA) können in Archaeobakterien, speziell in *Sulfolobales*, beispielsweise mit Oligonukleotiden der SEQ ID Nrn. 1 bis 4 identifiziert werden.

Ein Vektor, welcher eine für die erfindungsgemäße Ligase kodierende Nukleinsäure enthält, ist beispielsweise pDam-L3. Dieser Vektor wurde erhalten durch Klonierung eines 857 bp HindIII/EcoRI DNA-Fragments von *D. ambivalens* in pUC18 (host strain *E. coli* JM 83, Klon pL 8/2). Dazu wurde genomische DNA endonukleolytisch verdaut und Southern-Transfer und Hybridisierung nach Standardvorschrift (Maniatis et al., *Molecular Cloning, a laboratory manual*, Cold Spring Harbour Laboratory Press (1989) und anschließende Identifizierung des Fragments mit ³²P-markiertem Oligonukleotid SEQ ID 1 durchgeführt. Die Oligonukleotide SEQ ID 2 und 3 wurden benutzt um zwei verschiedene PstI/XbaI Fragmente zu identifizieren und in pUC18 (Plasmide pL1/2 und pL2/4) zu klonieren, welche den N-terminalen und den C-terminalen Teil des DNA-Ligase-Gens von *Desulfurolobus ambivalens* enthalten (FIG 2). Danach wurde pDam-L3 aus diesen drei Klonen durch Ligation des EcoRI/XbaI Fragments von pL 2/4 in p-BluescriptIIKS⁻ (Stratagene), gefolgt vom EcoRI-/HindIII-Fragment von pL 8/2 und dem HindIII Fragment von pL 1/2 zusammengesetzt (FIG 3).

Um das Gen in *E. coli* exprimieren zu können, wurde die codierende Sequenz mit Hilfe der Oligonukleotide SEQ ID NO 9 und SEQ ID NO 10 als Primer in einem PCR-Ansatz amplifiziert. Diese Primer enthalten im 5'-Bereich zusätzlich eine für das Restriktionsenzym BamHI-spezifische Sequenz. Nach BamHI-Verdau wurde das PCR-Fragment in den BamHI-gespaltenen Expressionsvektor pQE-8 kloniert (Diagen, QIAexpress Vektor Kit SRF, Literatur: *The QIAexpressionist*; Stüber et al., 1990, in *Immunological Methods*, eds. I. Lefkovits and B. Pernis, Vol. IV, 121-152). Das neukombinierte Plasmid wurde pQE-L6 bezeichnet. Nach Transformation in den *E. coli*-Stamm M15 wurde das Expressionsprodukt wie in Beispiel 8 gereinigt. Dabei enthält M15 zusätzlich das Plasmid pREP4, welches für das Neomycinphosphotransferase-Gen (neo) und das lacI-Gen (= lac-Repressor) codiert. Das neo-Gen vermittelt dem Stamm eine Kanamycin-Resistenz (zusätzlich zu der Ampicillin-Resistenz, welche über pQE-L6 eingebracht wurde). Der mit dem Plasmid pQE-L6 transformierte Stamm M 15 wurde bei DSM (Deutsche Sammlung von Mikroorganismen und Zellkulturen) unter Nr. 7070 am 19.05.1992 hinterlegt.

Die erfindungsgemäßen Nukleinsäuren hybridisieren nur mit DNA aus *Sulfolobus acidocaldarius* und *Sulfolobus shibatae*. Hybridisierung mit genomischen Sequenzen der Ligasegene von Organismen außerhalb der Archaeobakterien fand auch unter relaxierten Bedingungen nicht statt. Dies bedeutet, daß das Ligase-Gen eine Nukleotidsequenz hat, die sich stark von der der bekannten Ligasegene unterscheidet und daß die Ligasegene aus Archaeobakterien und insbesondere von Sulfolobaceae, untereinander starke Ähnlichkeiten zeigen.

Das Plasmid pDam-L3 enthält drei offene Leserahmen, wovon die Sequenz von Basenpaar 908 bis 2710 (FIG 1) der für die Ligase kodierenden Sequenz entspricht. Die Sequenz kodiert für ein Protein von 67619 D (errechnet), welches 600 Aminosäuren enthält. Der GC-Gehalt der kodierenden Region für die erfindungsgemäße DNA-Ligase beträgt ca. 36 %, leicht oberhalb des Gesamt-GC-Gehalts von *Desulfurolobus ambivalens* DNA (31 %). In FIG 1 ist die Nukleotidsequenz der erfindungsgemäßen Nukleinsäure und die entsprechende Aminosäuresequenz angegeben.

Der Fachmann kann diese Nukleotidsequenz dahingehend abändern, daß sich keine Änderungen in der Aminosäuresequenz ergeben (Degenerierung des genetischen Codes). Es ist anzunehmen, daß von Änderungen bis zu einem bestimmten Grad in der Nukleotidsequenz, welche auch zu Aminosäureänderungen führen, keinen Einfluß auf die Aktivität der erfindungsgemäßen Ligase zu erwarten ist. Bevorzugt umfaßt die Erfindung somit auch Nukleinsäuren, welche zu mehr als 40 %, bevorzugt mehr als 70 %, besonders bevorzugt jedoch mehr als 90 % zu der Sequenz in FIG 1 bzw. der genannten Teilsequenzen komplementär sind. Auch in Bezug auf die Länge des Ligasegens ist anzunehmen, daß eine Abweichung von der Länge des Gens aus *Desulfurolobus ambivalens* in gewissem Umfang keine entscheidende Änderung der Funktion der Kodierung für die erfindungsgemäße Ligase bewirkt. Bevorzugt ist die Nukleinsäure mindestens 150 bp, insbesondere zwischen 1000 und 3000 bp, lang. Die Nukleotidsequenz enthält insbesondere ein Motiv, welches für eine ATP-Bindungsstelle der Aminosäuresequenz K,D,G,R kodiert.

Die erfindungsgemäße Nukleinsäure kann auch dahingehend definiert werden, daß ihr komplementärer Strang unter den Bedingungen 6xSSC, 50 % Formamid, Raumtemperatur oder auch stringenter, mit nicht zum Vektor gehörigen Sequenzen und somit *Desulfurolobus*-spezifischen Sequenzen der Nukleinsäure aus FIG 1, insbesondere dem Teil der Nukleotide Nr. 908 bis 2716, besonders bevorzugt 988-2457, aus pDam-L3 oder mit dem 3382 bp langen XbaI/HindIII-Fragment aus *Desulfurolobus ambivalens* hybridisiert.

Solche Nukleinsäuren sind beispielsweise die Oligonukleotide der SEQ ID NO 1 - 4 und 7 - 10. Sie können beispielsweise als Primer zur templatabhängigen Extension (z.B. PCR) oder nach Einführung einer Markierung als Nachweissonden für *Desulfurolobus ambivalens* bzw. des Ligasegens eingesetzt werden.

Ebenfalls Gegenstand der Erfindung ist ein Verfahren zum Nachweis von Nukleinsäuren unter Verwendung der erfindungsgemäßen Ligase zur Verknüpfung zweier Nukleinsäuren, welche in benachbarten Bereichen mit einer nachzuweisenden Nukleinsäure hybridisiert sind.

Ein Beispiel für einen Nachweis unter Verwendung einer Ligase ist in der EP-A-0 320 308 beschrieben. Die Bedingungen für die Ligase-Chain-Reaction können aus diesem Dokument übernommen werden, wenn für die dort beschriebene Ligase die erfindungsgemäße Ligase eingesetzt wird. Die ATP-Abhängigkeit der erfindungsgemäßen Ligase muß jedoch berücksichtigt werden. Aus diesem Grund sollte ATP während der Ligasereaktion in einer Konzentration von mehr als 0.1 mmol/l, bevorzugt zwischen 0.1 und 20 mmol/l vorhanden sein. Die Konzentration von Polyethylenglykol (PEG) kann bei dem erfindungsgemäßen Verfahren 0 bis 20 % bevorzugt 5 bis 10 % betragen.

Ein weiteres Verfahren zum Nachweis von Nukleinsäuren, bei dem die erfindungsgemäße Ligase zur Ligation von benachbarten Nukleinsäuren eingesetzt werden kann, ist in WO 90/01069 beschrieben. Hierbei wird zunächst eine Lücke von wenigen Nukleotiden zwischen benachbart hybridisierten Nukleotiden durch Einbau von Mononukleosidtriphosphaten aufgefüllt und die entstandene Lücke (ein sogenannter Nick) mit der Ligase geschlossen.

In FIG 1 (FIG 1a und FIG 1b) ist die Nukleotidsequenz von pDam-L3, inklusive der vollständigen Sequenz für die DNA-Ligase aus *Desulfurolobus ambivalens* und pDam-L3 ORF 3 angegeben, sowie Teile der Sequenzen von pDam-L3 ORF 2 und 4. Unterstrichen: Box A; Fett: Transkriptionsstart.

In FIG 2 ist das Klonierungs- und Sequenzierungsschema für Plasmid pDam-L3 gezeigt.

Die gestrichelten Pfeile markieren Sequenzen, die mit den Oligonukleotiden L1 (SEQ ID NO 2) und L2 (SEQ ID NO 3) gewonnen wurden und zur Überprüfung der Ligationsergebnisse dienten, durchgezogene Pfeile die übrigen Sequenzen, gewonnen durch Subklonierung der entsprechenden Fragmente. Die lang gestrichelte Linie markiert die Lage des DNA-Ligasegens, rechts und links gestrichelt ist die Fortsetzung der Plasmide pL-1/2 und 2/4 angezeigt.

FIG 3 enthält die Plasmidkarte von pDam-L3 mit Restriktionsschnittstellen. Ebenfalls eingezeichnet sind die Orientierungen der Leserahmen. Die 0-Marke entspricht der 0-Marke von p-Bluescript IKS[™]. lig entspricht dem DNA-Ligasegen.

FIG 4 enthält einen Vergleich der Aminosäuresequenzen bekannter DNA-Ligasen. Auf der rechten Seite ist die Aminosäurenumerierung der individuellen Sequenz angegeben. Über der Sequenz ist die Numerierung der längsten Ligasesequenz (*H. sapiens*) sowie eine besonders homologe Stelle (aa 587 bis 593) angegeben. Fettgedruckt sind hochkonservierte Teile. Der Pfeil zeigt die ATP-Bindungsstelle (Lysin). Ein Eintrag in der Konsensuslinie (letzte Zeile, Kons) wurde im Fall von mehr als 50 % Identität gemacht. Ein Stern entspricht 1:1 Verteilung, Großbuchstaben bedeuten Identität über alle verglichenen Sequenzen.

FIG 5 zeigt einen paarweisen Vergleich der bekannten DNA-Ligasesequenzen unter Einschluß der RNA-Ligase von Bacteriophagen T4 und der tRNA-Ligase von *S. cerevisiae*. Im oberen rechten Dreieck sind die paarweisen Identitäten (%), bestimmt mit dem Programm GAP, angegeben, im unteren linken Dreieck die Signifikanzanalyse der Sequenzhomologie mit dem Programm RDF 2. Alle Werte der RDF 2-Bestimmungen sind SD-Einheiten, daß heißt die Trefferquote der paarweisen Ausrichtung wurden durch die Standardabweichungen der zufälligen Ausrichtungen geteilt. In der mittleren Diagonale sind zur Kontrolle die RDF 2-Ergebnisse der Sequenzen mit sich selber aufgetragen. In Klammern ist die Länge der jeweiligen Sequenz in Aminosäuren angegeben. In beiden Fällen wurde die jeweils längere Sequenz auf die Länge der kürzeren Sequenz reduziert, um Zufallsübereinstimmungen auszuschließen.

In FIG 6 sind die Ergebnisse der Transkriptionsstudien an Gesamt-RNA aus *D. ambivalens* gezeigt. Fette Pfeile: Hauptstart- und Terminationssignale. Unterstrichen: Box A und Box B Motive.

In FIG 7 ist die Plasmidkarte von pQE-L6 gezeigt. Charakteristische Restriktionsstellen für das Ligasegen sind angegeben. Die Details der Vektorkonstruktion sind zu entnehmen aus "QIA-Expressionist (DIAGEN GmbH, Düsseldorf, BRD)".

Die Erfindung wird durch die folgende Beispiele weiter erläutert:

Beispiel 1

Identifizierung des DNA-Ligase-Gens in *Desulfurolobus ambivalens*

Chromosomale DNA von *D. ambivalens* wurde aus anaerob gezogenen Zellen nach gängigen Methoden präpariert (Zellaufschluß mit SDS, Phenol- und Chloroformextraktionen, Ethanol-fällung, zweimalige CsCl-Dichtegradientenzentrifugation (Molecular Cloning; A laboratory manual CSH Laboratory Press 1989). Gesamt-RNA aus anaerob und aerob gezogenen Zellen wurde mittels der Guanidinium-Thiocyanat-Methode präpariert (Anal. Biochem. 162, 156-159 (1987)). Die Präparation von Plasmiden erfolgte nach Angaben des Herstellers über Quiagen-Säulen geeigneter Kapazität.

Oligodesoxynukleotide SEQ ID NO 1-4 und 7-10 wurden auf einem DNA Synthetisator 381 (Applied Biosystems, Weiterstadt) synthetisiert.

Die Oligonukleotide wurden mit ^{32}P -ATP endmarkiert, indem sie in PNK-Puffer (J. Mol. Evol. 24, 167-173 (1986)) im Verhältnis 1 pmol Oligonukleotid zu 10 μCi mit 5 U T4-Polynukleotidkinase für 30 min bei Raumtemperatur inkubiert wurden. Die Prähybridisierung von Filtern erfolgte in 6 x SSC, 1 x Denhardt's 0,5 % SDS, 20 $\mu\text{g/ml}$ tRNA (*E. coli*) und 0,05 % tetra-Natriumdiphosphat-10-hydrat bei 65 °C für 30-120 min. Die Hybridisierung erfolgte in dem oben genannten Puffer mit dem Unterschied, daß nur 0,1 % SDS zugegeben wurde. Hybridisierungstemperatur und -zeit wurden empirisch ermittelt. In der Regel wurden die Filter nach der Hybridisierung einmal für 15 min bei Raumtemperatur in 1 x SSC/0,5 % SDS gewaschen und einmal für 30 min bei der als optimal ermittelten Hybridisierungstemperatur. Koloniehybridisierungen (Molecular Cloning loc. cit.) wurden grundsätzlich genauso behandelt.

Beispiel 2

Klonierung

Ein 857 bp großes HindIII/EcoRI-Fragment wurde in pUC 18 (Amersham, Braunschweig) (pL 8/2) kloniert, nachdem es durch Spaltung genomischer DNA, Southern-Transfer und Hybridisierung mit dem ^{32}P -markierten Oligonukleotid Sox1 (SEQ ID NO 1) identifiziert worden war.

Von der Nukleinsäuresequenz von pL-8/2 wurden zwei weitere Oligonukleotide abgeleitet, mit deren Hilfe zwei PstI/XbaI-Fragmente identifiziert und kloniert wurden, die das N-terminale (pL-1/2, Oligonukleotid SEQ ID NO 2) und das C-terminale (pL-2/4, Oligonukleotid SEQ ID NO 3) Ende des Genes enthielten. Letzteres enthielt offenbar aufgrund einer Doppelligation oder partieller Spaltung noch eine interne XbaI-Schnittstelle (FIG 2). Ausgehend von den drei teilweise überlappenden Klonen wurde das Plasmid pDam-L3 zusammengesetzt, das die vollständige genetische Information für die DNA-Ligase trug. Dazu wurden das EcoRI/XbaI-Fragment von pL2/4 zunächst in pBluescript II KS⁻ (Stratagene Cloning Systems La Jolla, USA)

kloniert, gefolgt von dem EcoRI/HindIII-Fragment von pL8/2 und dem HindIII-Fragment von pL1/2. Die Ergebnisse der einzelnen Klonierungsschritte wurden jeweils durch Doppelstrangsequenzierung mit den Oligonukleotiden SEQ ID NO 2, 3, 13 und 14 überprüft (vgl. FIG 2).

- Das resultierende Plasmid pDam-L3 enthielt ein DNA-Fragment von 3382 bp Länge mit der vollständigen Information für die DNA-Ligase von *D. ambivalens* und des weiteren ca. 900 bp stromauf und ca. 700 bp stromab (FIG 3).

Beispiel 3

10 Subklonierung und Sequenzierung von pDam-L3

- Geeignete Restriktionsfragmente wurden in die Vektoren M13mp10, M13tg131, M13mp18, M13mp19 (Amersham, Braunschweig BRD), bzw. in die vier pBluescript-Vektoren (Stratagene Cloning Systems, La Jolla, USA) kloniert. Die Sequenzierungsstrategie ist in FIG 2 dargestellt. Sequenzreaktionen wurden je nach Erfordernissen mit Einzel- oder Doppelstrangmatrizen gemacht. Sequenzreaktionen mit in M13-Vektoren subklonierten Fragmenten wurden mit einem Universal-Oligonukleotid gestartet (= -40-Oligo des Sequenase Kits, United States Biochemical Corporation, Cleveland, Ohio, USA, SEQ ID NO 11), Reaktionen in pUC mit dem Universal bzw. dem reversen - 40 Oligonukleotid (für das reverse Oligonukleotid wurden die Angaben in Nucl. Acids Res. 17, 4897 (1989) berücksichtigt), diejenigen in pBluescript-Vektoren je nach Erfordernissen mit den Oligonukleotiden SEQ ID NO 13 und 14. Bei letzteren wurde je nach Orientierung des Fragmentes im Vektor der Einzelstrang nach Infektion mit dem Helferphagen VPS13 (Stratagene Cloning Systems) für die Sequenzierung präpariert, oder aber das Fragment Doppelstrang-sequenziert. Zusätzlich wurden Doppelstrang-Sequenzierungen mit spezifisch synthetisierten Oligonukleotiden oder den Sucholigonukleotiden gemacht (s.o.). Für die Sequenzreaktionen wurde der Sequenase-Kit entsprechend den Anleitungen des Herstellers verwendet (United States Biochemical Corporation). Sequenzgele, gegossen mit Sequagel-Lösungen (National Diagnostics Corporation, Manville, New Jersey, USA), wurden in Makrophor-Sequenzierkammern gefahren (Pharmacia, LKB, Freiburg).

Ergebnisse

- Die vollständige Nukleotidsequenz und ihre Übersetzung in die Aminosäuresequenzen sind in FIG 1 dargestellt. Vier offene Leserahmen innerhalb des sequenzierten Bereiches wurden näher analysiert. Der längste (bp 908-2710, FIG 2) enthielt die Information für ein hypothetisches 67617 Da-Protein bestehend aus 600 Aminosäuren. Von den drei weiteren analysierten Leserahmen begann einer bei bp 2961 und ging über das Ende der sequenzierten Region hinaus (pDam-L3-ORF 2), der zweite überspannte bp 771-250 (pDam-L3-ORF 3, vgl. FIG 3), der Anfang des dritten lag im nicht sequenzierten Bereich von pL1/2 und endete bei bp 273 (pDam-L3-ORF4).

Beispiel 4

40 Alignment mit pro-/eukaryonten DNA/RNA-Ligasen

- Die erhaltenen Nukleotidsequenzen wurden mit Hilfe des Programmpaketes UWGCG, (Nucl. Acids Res. 12, 387-395 (1984)) Version 7.0 (1991) analysiert. Vergleiche von offenen Leserahmen mit in der Genbank Mipsx (Martinsried) gespeicherten Sequenzen wurden mit Hilfe des Programmes PIRFASTA (Meth. Enzymol. 183, 63-99 (1990)) angestellt. Die verfügbaren Aminosäuresequenzen der DNA-Ligasen von *Schizosaccharomyces pombe* (Enz. J. Biochem. 162, 659 (1987)), *Saccharomyces cerevisiae* (Nucl. Acids. Res. 13 8323-8337 (1985)), *Escherichia coli* (Mol. Gen. Genet. 204, 1-7 (1986)), *Thermus thermophilus* (J. Bacteriol. 173, 5047 (1991)), der menschlichen DNA-Ligase I (Proc. Natl. Acad. Sci. USA 87, 6679 (1990)), der Vaccina Virus DNA-Ligase Nucl. Acids Res. 17, 9051-9062 (1989), der Ligasen der *E. coli*-Bakteriophagen T3, J. Mol. Biol. 479-495 (1987), T4, Nucl. Acids Res. 11, 7145-7156 (1983), T6, Dokl. Akad. Nank SSSR 299, 737-742 (1988) und T7, J. Mol. Biol. 148, 303-330 (1981) der RNA-Ligase von T4, EMBO J. 3, 397-402 (1988) und der tRNA-Ligase von *S. cerevisiae*, J. Biol. Chem. 263, 3171-3176 (1988) wurden Sequenzvergleichen mit dem Programm GAP unterworfen, um die paarweisen Identitäten zu bestimmen (FIG 5). Nicht berücksichtigt wurden die Aminoacyl-tRNA-Ligasen, die eine ganz andere Funktion ausüben. Die Signifikanzanalyse der paarweisen Ausrichtungen wurde mit dem Programm RDF2 durchgeführt. Dieses Programm vergleicht die Ergebnisse von paarweisen Ausrichtungen mit den Ergebnissen einer einzugebenden (hier: 1000) Anzahl von Ausrichtungen, bei dem die eine der beiden Sequenzen in zufälliger Weise

aufgebrochen und gemischt wird (= "shuffeln"), und es berechnet die Standardabweichung daraus. In FIG 5 sind die Vielfachen des Ergebnis der Ausrichtung geteilt durch die Standardabweichungen angegeben (SD-Einheiten). Es wurden Tests in beiden Richtungen durchgeführt, d.h. von zwei Sequenzen wurden beide jeweils einmal intakt gelassen und "geschuffelt" und die beiden Ergebnisse gemittelt. Da die ATP-abhängigen DNA-Ligasen am C-terminalen Ende stärker konserviert waren, wurden die jeweils längeren Sequenzen von N-Terminus her auf die gleiche Länge gekürzt.

Die sich aus den paarweisen Vergleichen ergebenden Prozentwerte der Identitäten, bzw. der Vielfachen der Standardabweichungen von paarweiser und zufälliger Ausrichtung (RDF2) der Aminosäuresequenzen der DNA-Ligasen sind in FIG 5 dargestellt. Auffallend ist hier, daß die DNA-Ligase von *D. ambivalens* bei allen Unterschieden mehr als 30 % Sequenzidentität zu allen vier bekannten ATP-abhängigen DNA-Ligasen der Eukarya hat und Werte zwischen dem 48- und 86-fachen der Standardabweichung in der RDF2 aufweist (FIG 5). Die RDF2-Ergebnisse der einzelnen Sequenzen, kontrollhalber gegen sich selber getestet (100 % Identität) lagen im Bereich zwischen 200 und 450 SD und waren vorwiegend von der Länge der Sequenz abhängig. Keine signifikante Ähnlichkeit wurde zu den NAD-abhängigen, bakteriellen, cytoplasmatischen DNA-Ligasen von *E. coli* und *Th. thermophilus* gefunden (<20,4 % Sequenzidentität; < 0,2 SD in der RDF2). Ebenso wenig war eine Ähnlichkeit zu der ebenfalls ATP-abhängigen RNA-Ligase vom *E. coli*-Phagen T4 und der tRNA-Ligase von *S. cerevisiae* feststellbar. Für die T-Phagen-Ligasen lagen die Ergebnisse dazwischen. Die RDF2-Analyse ergab zwar keine signifikanten Resultate, die Identitätswerte lagen jedoch bei 20,2-23,7 %. Betrachtet man FIG 5, so fallen folgende Punkte auf: Die eukaryalen und die *D. ambivalens*-DNA-Ligasen bilden untereinander eng verwandte Gruppe. Die T4- und T6-Ligasen sind bis auf sechs Positionen identisch. Die T3- und T7-Ligasen sind mit 72 % Identität ebenfalls eng verwandt, ebenso die NAD-abhängigen DNA-Ligasen von *E. coli* und *Thermus thermophilus* (46,7 %), dieses Paar ist mit dem Rest jedoch nicht verwandt. Die RNA-Ligasen sind weder untereinander noch mit dem Rest in Beziehung zu bringen. Die vier bekannten DNA-Ligasen der T-Phagen sind von den übrigen ATP-abhängigen DNA-Ligasen deutlich getrennt. In der paarweisen Analyse liegen sie mit Identitätswerten von 20,2-26,0 % über dem Rest, in der Signifikanzanalyse weisen sie jedoch Werte um den Nullpunkt auf. Sie können aber aufgrund konservierter Signaturen in die Ausrichtung der ATP-abhängigen DNA-Ligasen mit aufgenommen werden (s.u.). Die Vaccina-DNA-Ligase weist eine Zwischenstellung auf. Sie ist den übrigen eukaryalen Enzymen nahe verwandt, erzielte aber auch Werte von 4,3-8,3 in der RDF2-Analyse mit den T-Phagen-Ligasen, sie verbindet somit die beiden Gruppen von Ligasen miteinander.

Im folgenden werden nur die ATP-abhängigen DNA-Ligasen behandelt, da nur bei ihnen eine gemeinsame Ausrichtung möglich war. Für alle anderen konnten keine gemeinsamen Strukturen im Sequenzvergleich gefunden werden, die sich von Zufallstreffern abheben. Die DNA-Ligasen unterscheiden sich in ihrer Gesamtlänge erheblich (346-919 Aminosäurereste), wobei die eukaryalen generell länger als die T-Phagen-Ligasen sind (346-487 gegenüber 552-919 Aminosäurereste). Die Ausrichtungen der Proteinsequenzen der eukaryalen DNA-Ligasen untereinander zeigt viele konservierte Bereiche, die teilweise von Bereichen geringer Ähnlichkeit getrennt sind. Es war bereits in früheren Untersuchungen aufgefallen, daß speziell im N-terminalen Bereich die Ligasen aus den Hefen und dem Menschen wenig konserviert sind, und man hatte daraus den Schluß gezogen, daß diese Teile der Sequenzen keinerlei Beziehungen zueinander hätten. Drei hoch konservierte Strukturen wurden bei allen wiedergefunden, die z. T. schon in früheren Untersuchungen aufgefallen waren. Es handelte sich zum einen um das ATP-bindende Zentrum, das allerdings nur bei der Rinderligase I experimentell bestimmt worden war. Nahe bei der beschriebenen Konsensussequenz KYDGXR (K ist ATP-bindend, Pos. 650) wurden noch zwei wahrscheinlich ebenfalls zum Zentrum gehörende Strukturen gefunden (Pos. 560: PMLA, bzw. PFKA bei T3 und T7; Pos. 610: SR). Zum zweiten wurde die von Barker et al. (1985) aufgrund ihrer Konservierung fälschlicherweise als ATP-bindendes Zentrum vermutete Struktur bei Pos. 785-813 in der Ausrichtung wiedergefunden. Die dritte ist eine nahe dem C-Terminus liegende Struktur bei Pos. 945-960. Darüberhinaus erlaubten eine Reihe weiterer, weniger konservierter Strukturen die Ausrichtung der neun ATP-abhängigen DNA-Ligasen trotz geringer Signifikanz der Ähnlichkeiten zwischen T-Phagen und eukaryalen Ligasen. Das ATP-bindende Zentrum ist bei allen ATP-abhängigen DNA-Ligasen so hoch konserviert, daß die Zuordnung aufgrund der Sequenzhomologie zuverlässig erscheint.

Beispiel 5

55 Transkript-Analyse

Die Initiation der Transkription der offenen Leserahmen in vivo wurde mittels Primerextensionsanalyse bestimmt (Proc. Natl. Acad. Sci. USA 87, 5851 (1990)). Die in-vivo Termination des DNA-Ligase-Genes

wurde durch S1-Kartierung mit einem α -³²P-ATP markiertem NcoI/SnaBI-Fragment (bp 2458-3221) als Sonde bestimmt. Gesamt-RNA aus aerob und anaerob gewachsenen Zellen von *D. ambivalens* wurde für Transkriptionsanalyse des Ligasegens eingesetzt. Mit einer Ausnahme (RNA aus stationären Zellen) waren die Signale in der Primerextensionsreaktion mit Gesamt-RNA sehr schwach, es war aber aerob wie anaerob in gleicher Stärke nachweisbar. Die Transkription startete auf der ersten Base des ATG-Startcodons der Translation, bzw. auf dem vorgelagerten Thymidinrest. Von den übrigen Leserahmen von pDam-L3 wurde kein Startsignal erhalten.

Die Ergebnisse der Primerextensionsanalyse zeigten eindeutig, daß das DNA-Ligasegen in vivo transkribiert wird, obwohl die Stärke des Signals in der Regel gering und nur in einem Falle mit RNA aus stationären Zellen stark war. Die Promotersequenz wies die für die Archaea typische Struktur mit Box A und Box B auf. Über die übrigen Leserahmen des Fragmentes konnten keine Aussagen getroffen werden. Die Tatsache, daß kein mRNA-Startsignal gefunden wurde, schließt nicht aus, daß sie möglicherweise in bestimmten Fällen exprimiert werden könnten.

Die Termination der Transkription wurde durch S1-Kartierung an derselben RNA-Präparation bestimmt, die auch das starke Initiationssignal lieferte. Die Sequenz, der die Terminationssignale zugeordnet werden konnten, lautet ACATTATG, die vorgelagerte Basensequenz lautet AGTAGATGG. Die zugehörige Sequenz und die Lage des Signals sind in FIG 6 dargestellt. Die aus den Start- und Terminationssignalen abgeleitete Länge des Transkriptes betrug 1908-1914 Nukleotide.

Beispiel 6

Kreuzhybridisierung mit anderen Spezies

Genomische DNA aus den folgenden Organismen wurde jeweils doppelt mit BglII und NcoI verdaut, sowie zusätzlich mit dem den in Klammern angegebenen Enzym(en): Archaea: *Sulfolobus acidocaldarius* (EcoRI), *S. shibatae* (EcoRI), *Desulfurococcus mucosus* (EcoRI), *Thermoproteus tenax* (PstI), *Thermococcus celer* (PstI), *Thermoplasma acidophilum* (EcoRI), *Methanothermus fervidus* (EcoRI), *Methanococcus vannielii* (EcoRI) und *Haloferax volcani* (Bam HI/PstI); Bacteria: *Escherichia coli* (HindIII), und *Bacillus stearothermophilus* (EcoRI); Eucarya: *Box primigenius taurus* (Rind, EcoRI), und *Saccharomyces cerevisiae* (Bäckerhefe, EcoRI). Der Auftrennung der Ansätze auf Agarosegelen folgten ein Southern-Transfer und die Hybridisierung der Filter mit einem radioaktiv markierten BglII/NcoI-Fragment aus pDam-L3 (bp 988-2457, FIG 3), welches ca. 82 % des Genes enthält. Die Bedingungen waren: 6 x SSC, 50 % Formamid, Raumtemperatur; die Waschbedingungen waren 2 x SSC, Raumtemperatur. Mit in die Auftrennung eingeschlossen wurde genomische DNA von *D. ambivalens* (BglII/NcoI und EcoRI-geschnitten), sowie das ebenfalls BglII/NcoI-geschnittene Plasmid pDam-L3.

Die Ergebnisse der Hybridisierung zeigen, daß *D. ambivalens* offenbar nur eine Kopie dieses Genes trägt. Außerdem konnten definierte Banden in der DNA von *Sulfolobus acidocaldarius* und *S. shibatae* nachgewiesen werden, zwei Spezies, die innerhalb der Ordnung der Sulfolobales die am fernsten verwandten Organismen sind. Bei allen anderen Organismen konnte keine Hybridisierung erzielt werden. Aus diesen Ergebnissen läßt sich ableiten, daß innerhalb der Sulfolobales offenbar das Gen hoch konserviert ist. Darüber hinaus konnte mit dieser Methode auch keine enge Verwandtschaft zu den eukaryalen Enzymen auf DNA-Ebene festgestellt werden.

Zum jetzigen Zeitpunkt umfaßt die Ordnung der Sulfolobales die folgenden Spezies:

Sulfolobus acidocaldarius

S. solfataricus

S. shibatae

S. thuringiensis

S. metallicus

Acidianus infernus

A. brierleyi

Desulfurolobus ambivalens

Stygiolobus azoricus

Metallosphaera sedula

Weitere Neuisolat sind zu erwarten. Alle diese Spezies haben eine vergleichbare Morphologie und Lebensweise und sind auch phylogenetisch verwandt.

Beispiel 7Expressionsklonierung

Konstruktion des Expressionsplasmids pQE-L6: Für die Expression des DNA-Ligase-Genes in *E. coli* wurde aus einer Präparation genomischer DNA mit Hilfe der Polymerase-Kettenreaktion der kodierende Bereich des Genes amplifiziert. Die Oligonukleotide SEQ ID NO 9 und 10 wurden so gestaltet, daß N-terminal eine BamHI-Schnittstelle gefolgt von einer Schnittstelle der Restriktionsprotease Faktor Xa eingefügt worden war und am C-Terminus eine BamHI-Schnittstelle. Die Amplifikationsbedingungen waren:

1 x Reaktionspuffer (Boehringer Mannheim, Kat. Nr. 1146165)
 je 100 µM dNTP
 2,5 U Taq-DNA-Polymerase
 je 40 pmol Oligos
 500 ng genom. DNA

| | |
|-----------------|--------------|
| Denaturierung: | 94 °C, 60 s |
| Anlagerung: | 51 °C, 120 s |
| Polymerisation: | 72 °C, 120 s |
| 40 Zyklen | |

Das Reaktionsprodukt wurde nach Spaltung mit BamHI in den ebenfalls mit BamHI geschnittenen und mit alkalischer Phosphatase dephosphorylierten Vektor Quia Express pQE-8 (DIAGEN GmbH, Düsseldorf, BRD) ligiert.

Expression in *E. coli*: Nach Transformation von *E. coli* wurde eine Vorkultur über Nacht auf Kanamycin- und Ampicillin-haltigem 2 x TY-Medium angezogen. Die daraus angeimpfte Hauptkultur im gleichen Medium wurde bis zu einer Zelldichte von $OD_{600} = 0,8-1,0$ gezogen und dann mit einer IPTG-Konzentration von 2 mM induziert. Der zeitliche Verlauf der Proteininduktion wurde über SDS-Gele verfolgt; in der Regel wurden die Zellen 3 h nach Inkubation geerntet. Die Zellen wurden nach Vorbehandlung mit Lysozym aufgeschlossen, indem sie dreimal in flüssigem Stickstoff eingefroren und wieder aufgetaut wurden. Die Hitzefällung von *E. coli*-Proteinen erfolgte durch fünfzehnminütige Inkubation bei 75 °C und anschließender Zentrifugation für 2 min in der Eppendorfzentrifuge.

Die PC-Reaktion mit Oligonukleotiden, die eine Amplifikation des DNA-Ligasegenes unter Einfügung einer Faktor Xa-Schnittstelle im Protein erlaubten (IEGR), ergab eine einzige scharfe Bande von ca. 1900 bp im Agarosegel. Sie wurde in pQE-8 kloniert, so daß das Oligopeptid der Sequenz MRGSHHHHHGSIEGR N-terminal mit der DNA-Ligase fusioniert wurde. Die [6 x His]-Sequenz soll dabei eine Affinitäts-Reinigung des Fusionsproteins über Nitrilotriessigsäure-cheliertes Nickel ermöglichen. Identifizierung durch Koloniehybridisierung wurden fünf verschiedene positive Klone angezogen, die erhaltenen Plasmide durch Restriktionsspaltung auf das Vorhandensein eines Fragmentes der erwarteten Größe untersucht und der Verlauf der Expression nach IPTG-Induktion im SDS-Gel verfolgt. Vier von den fünf Klonen zeigten eine zusätzliche Protebande mit einem apparenten Molekulargewicht von 65000 im SDS-Gel. Das erwartete Molekulargewicht des Fusionsproteins betrug 69487. Durch eine Hitzebehandlung von Rohlysaten der induzierten Zellen konnte ein großer Teil des Zellproteins abgetrennt werden, während das induzierte Protein löslich blieb. Bei dem fünften Klon wie auch bei der Vektorkontrolle fehlte die Bande. Es wurde vermutet, daß der fünfte, in der Hybridisierung positive Klon das amplifizierte Fragment in inverser Richtung enthielt.

Beispiel 8Reinigung der *D. amibvalens*-Ligase

E. coli-Bakterien, welche das Plasmid pQE-L6 enthalten, z.B. M15-pQE-L6, werden in 2 l DYT-Medium (J. Sambrook et al., Molecular Cloning, 1989, CSH, Vol. 3, Seite A.3) mit 20 µg/ml Ampicillin und 20 µg/ml Kanamycin bei 37 °C angezüchtet. Bei $OD_{550} = 0,8$ wird mit 1 mM IPTG induziert und nach 3 Std.- bis Übernacht-Inkubation bei 37 °C geerntet. Die Bakterienzellen werden bei 6000 rpm in 500 ml - Bechern im GSA-Rotor einer Sorvall-Zentrifuge pelletiert und in 20 ml Tris, 20 mM, pH 7.6; EDTA, 1 mM und β -Mercaptoäthanol, 10 mM resuspendiert. Die Bakterienzellen werden mit Ultraschall-Behandlung aufgeschlossen. Dazu werden 4 x 5 ml Aliquots 5 x 1 Minute bei 50 % der Ausgangsleistung eines Branson Sonifier 450 sonifiziert. Danach wird die Suspension bei 39.000 g in einem SS34-Rotor in einer Sorvall-

Zentrifuge zentrifugiert und der Überstand mit 2 M KCl auf eine Salz-Endkonzentration von 300 mM gebracht.

Um die bakterielle DNA aus dem Ansatz zu reinigen, wird die Lösung auf eine DEAE-Säule (5 ml, equilibriert mit 300 mM KCl) gebracht und mit 70 ml 300 mM KCl eluiert. Der Säulendurchfluß wird 20' bei 65 °C behandelt, um die E. coli Proteine zu denaturieren. Anschließend wird im SS34-Rotor bei 18500 rpm 15 Minuten zentrifugiert und der Überstand mit gesättigtem Ammoniumsulfat (Volumenverhältnis 1 : 1) gefällt. Anschließend wird erneut bei 18000 rpm zentrifugiert und das Pellet, welches die Ligase aus *D. sulfurolobus* enthält, in 4 ml H₂O aufgenommen.

10 Beispiel 9

Ligationsaktivität der *Desulfurolobus ambivalens* - Ligase

Als Target für diesen Aktivitätsnachweis wird genomische DNA des Phagen mit den Restriktionsenzymen SmaI (blunt end) und Sall (2 Nukleotide 5'-überstehende Enden) geschnitten. Bei hoher Reaktionstemperatur sind Hybride dieser Enden nicht stabil, so daß nur die beiden Endfragmente des Genoms, welche kohäsive 5'-überstehende Enden von 12 nt besitzen (sog. cos sites) als Hybrid vorliegen. Dabei liegen in diesem Hybrid zwei sogenannte "nicks" in 12 nt Abstand vor. An diesem Template kann demnach die Aktivität einer Ligase an genickter DNA getestet werden: als Produkt entsteht ein Fragment aus den beiden

eine cos-site enthaltenden genomischen Endfragmenten.
3 µl der im Beispiel 8 beschriebenen *Desulfurolobus ambivalens* Ligase-Präparation werden in einem Gesamtvolumen von 50 µl in Tris HCl, 20 mM, pH 7.5; MgCl₂, 10 mM; Dithiothreitol (DTT), 10 mM; ATP, 1 mM; Polyethylenglycol (PEG), 5% bei 56 °C zwischen 1 Stunde und 12 Stunden inkubiert. Die Reaktionsprodukte werden in einem 1 % Agarosegel aufgetrennt und mit Ethidiumbromid-Färbung sichtbar gemacht.

Vorgeschlagene Variationsbreiten für Ligation:

| | |
|---------------------|---|
| Volumen: | 10 - 200 µl |
| Puffer: | 10 - 50 mM |
| MgCl ₂ : | 2 - 20 mM |
| DTT: | 0 - 20 mM |
| ATP: | 0.1 mM - 20 mM |
| PEG: | 0 - 20% |
| Temperatur: | Raumtemperatur bis 85 °C, Reaktionszeit 10 Min. bis Übernacht-Inkubation. |

Wird zu der beschriebenen Ligasereaktion zusätzlich Adenosin-5'-O-(3-thiotriphosphat) (ATP-Gamma-S) 10 mM bzw. 100 mM zugegeben, ist keine Ligation der beiden Lambda-Endfragmente zu sehen. Dies weist auf eine deutliche Kompetition des Co-Faktors ATP hin und beweist die ATP-Abhängigkeit der hier verwendeten Ligase.

40 Beispiel 10

Vergleichende Cofaktorbestimmung

Im vorliegenden Beispiel wird ein Versuch beschrieben, mit welchem die ATP-Abhängigkeit der erfindungsgemäßen Ligase und die Unabhängigkeit von NAD ermittelt wurde. Zum Vergleich wurden die Cofaktoren von aus dem Stand der Technik bekannten Ligasen (T4-Ligase, Firma: Boehringer Mannheim GmbH) und *Thermus spec.*-Ligase (Ampligase, Firma: Biozym) ermittelt.

Zur Cofaktorbestimmung diente die Adenylierungsreaktion. Hierbei wird radioaktiv markierter Cofaktor (³²PATP oder ³²PNAD) in einem Ligationspuffer mit der Ligase inkubiert. Dadurch wird die Radioaktivität durch die Phosphorylierung der Ligase (kovalent) auf das Protein übertragen. Nach Elektrophorese in einem SDS-Gel ist die Ligase durch Autoradiographie sichtbar.

Verwendet wurden der T4-Ligase-Puffer sowie der Ampligase-Puffer.

Die Inkubationszeit betrug jeweils 1 Stunde.

Die Inkubationstemperatur betrug für die T4-Ligase 37 °C, für die Ampligase 65 °C und für die *Desulfurolobus* Ligase 45 °C und 65 °C. Verwendet wurden von der T4-Ligase 1 U, von der Ampligase 100 U und von der *Desulfurolobus*-Ligase 10 bis 1000 ng pro Ansatz. Das Ansatzvolumen betrug 20 µl. Bei höheren Temperaturen wurden die Ansätze mit Öl überschichtet.

T4-Ligase-Puffer: 66 mM Tris/HCl, 5 mM MgCl₂, 1 mM DTT, pH 7,5 bei 20 °C

Ampligase-Puffer: 20 mM Tris/HCl, 50 mM KCl, 10 mM MgCl₂, 1 mM EDTA, 10 mM DTT
pH 7,6 bei 65 °C

Radioaktivität: Amersham

5 ³²PATP, spez. Aktivität: 15 TBq/mmol, 0,37 MBq Pro Ansatz

³²PNAD, spez. Aktivität: 37 TBq/mmol, 0,37 MBq pro Ansatz

Die Ampligase zeigte wie erwartet die NAD-Abhängigkeit. Die Desulfurolobus-Ligase wird eindeutig durch ATP adenyliert. Sie zeigt auch ein schwaches Signal bei NAD als Cofaktor. Dies ist aber auf eine Verunreinigung des NAD mit ATP zurückzuführen, da auch die T4-Ligase, als eindeutig ATP-abhängiges

10 Enzym, mit NAD ein schwaches Signal erzeugt.

Der vorstehende Versuch zeigt eindeutig, daß die erfindungsgemäße Ligase im Gegensatz zu der aus einem thermophilen Eubakterium bekannten Ligase ATP-abhängig ist.

SEQUENZPROTOKOLL

(1) ALLGEMEINE INFORMATION:

(i) ANMELDER:

(A) NAME: Boehringer Mannheim GmbH

25 (B) STRASSE: Sandhoferstr. 116

(C) ORT: Mannheim

(E) LAND: DE

30 (F) POSTLEITZAHL: 6800

(G) TELEPHON: 0621 759 4348

35 (ii) ANMELDETITEL: Thermostabile Ligase aus Archaeobakterien

(iii) ANZAHL DER SEQUENZEN: 14

(iv) COMPUTER-LESBARE FORM:

(A) DATENTRÄGER: Floppy disk

(B) COMPUTER: IBM PC compatible

45 (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)

(vi) FRÜHERE ANMELDEDATEN:

50 (A) ANMELDENUMMER: DE 42 17 134.2

(B) ANMELDEDATUM: 23-MAY-1992

(2) INFORMATION ZU SEQ ID NO: 1:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 36 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

TATGTINGCNA TWAATATGGC NGAATTRA AAA AAYGAA

36

(2) INFORMATION ZU SEQ ID NO: 2:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 20 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

TCCAGTGCTT GTAGCAACTT

20

(2) INFORMATION ZU SEQ ID NO: 3:

- 5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 10 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

15 (ii) ART DES MOLEKÜLS: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

20 GATTATGTAC ATATAGCCAC

20

(2) INFORMATION ZU SEQ ID NO: 4:

- 25 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 17 Basenpaare
 (B) ART: Nukleinsäure
 30 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

35 (ii) ART DES MOLEKÜLS: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

40 TAAGTTGAAG OCTAGAA

17

45

50

55

(2) INFORMATION ZU SEQ ID NO: 5:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 3382 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Doppel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: promoter
- (B) LAGE: 878..883

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: CDS
- (B) LAGE: 908..2707

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: terminator
- (B) LAGE: 2713..2722
- (D) SONSTIGE ANGABEN: /note= "potentiell"

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: terminator
- (B) LAGE: 2804..2811
- (D) SONSTIGE ANGABEN: /note= "experimentell"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

AAGCTTAAGA ATCTGACIT ATTCTTTGAA AAAATTGAAA AGCAAGGTGG GAAAGTTCTA 60

TATACTGATA AAGAAAGAAT AGCGTAGAT GTGATTTCT TTAATAATTT TGTAAAGAAA 120

CTTTCAGATA TTCACACTTC AGATGATGTA GAAGOGCAA AATACCTCAC TAAGCAAGAA 180
 5 CATAAATTAT TOCAAAAATT GAGACAATTG GGAATAATAT ACTTCGATAA TAATACTAAG 240
 GCATGGAAGT TATOGGAATT TAAATTTTGG TAAAGTTAOG CCTTGTGTGC CTGGGTATTT 300
 10 GCCATGATAA GGATTCTCTA CTGGTGTAA TGTCTGGOG AAAATGAGGT GAAGAAAGOG 360
 AGTCCCCCTC TTTATCTTCA OGGGAAAAGC AGAACCTAGT ACTTCTATTG TGAGTTGTCC 420
 15 TTCAAAGCCT GCATCTACGA TAGTOGGTGG TACAAAAGA CCCAGCCTGG CAAAAGAAGA 480
 CCTTAAGTTG ACGAAAGOCA TTACGTGTTT OGGCAGTTA ACATACTCCT CAGTAACTAA 540
 20 CAGTACATGT TOGTTAGGGT AAATGATGAA TTCATCTCCT TTCTCTATCT CATAAAATGA 600
 ACGGGGCTCT TTTCATCTT CATATATTTC GTCAGTCTTC TTAAAGGGG CTATCTCTCC 660
 25 TOCTACTCTT AAATGACCC CGTCTCTCT TATAGTATCT TGAGTAAGAG GAGAGATTAC 720
 TATCCAACC TTTTCTAGGT AGTATTTTAA ATCTCTATCT CCAAGAATCA TAAATATCTA 780
 CTATTTTTC TCTTGCAAGA ATATATTTTA TAAGGAAAT GGCTTCAGAG TAACTCTGAT 840
 35 ACTTCATCCT TCTGGCTTGC OGGTTTTCAT CATCAATAT ATTATATOGT AATTCAAAC 900
 TTTATTT ATG GAG TTT AAA ATT ATA GCT GAA TAC TTT GAT AGA CTA GAA 949
 40 Met Glu Phe Lys Ile Ile Ala Glu Tyr Phe Asp Arg Leu Glu
 1 5 10
 AAG ATC TCT TCT AGG CTT CAA CTT ACA TCA TTA CTT GCA GAT CTT TTT 997
 45 Lys Ile Ser Ser Arg Leu Gln Leu Thr Ser Leu Leu Ala Asp Leu Phe
 15 20 25 30

50

55

| | | |
|----|--|------|
| 5 | AAG AAA ACT GAC AAA AAT GTT ATT GAT AAA GTT GTA TAT ATA ATT CAG Lys Lys Thr Asp Lys Asn Val Ile Asp Lys Val Val Tyr Ile Ile Gln | 1045 |
| | 35 40 45 | |
| 10 | GGT AAA CTA TGG OCA GAC TTC TTA GGA ATG OCA GAA TTA GGA ATA GGT Gly Lys Leu Trp Pro Asp Phe Leu Gly Met Pro Glu Leu Gly Ile Gly | 1093 |
| | 50 55 60 | |
| 15 | GAA AAA TTC CTT ATA AGA GCA TTA TCA ATA GCT ACA AGC GTA AGC GAT Glu Lys Phe Leu Ile Arg Ala Leu Ser Ile Ala Thr Ser Val Ser Asp | 1141 |
| | 65 70 75 | |
| 20 | GAT GAA ATA GAA AAA ATG TAT AAA TCA GTA GGA GAT TTG GGG CAA GTA Asp Glu Ile Glu Lys Met Tyr Lys Ser Val Gly Asp Leu Gly Gln Val | 1189 |
| | 80 85 90 | |
| 25 | GCT TTT GAT ATA AAG CAA AAA CAG CAG TCT GGC AGT ATT TTA GCG TTT Ala Phe Asp Ile Lys Gln Lys Gln Gln Ser Ala Ser Ile Leu Ala Phe | 1237 |
| | 95 100 105 110 | |
| 30 | TTA GGA GCG CAA AAA GCT TCA AAG CCG CTA ACT GTA GAA AAA GTT TAT Leu Gly Ala Gln Lys Ala Ser Lys Pro Leu Thr Val Glu Lys Val Tyr | 1285 |
| | 115 120 125 | |
| 35 | GAT GAT TTA GCA AAA GTT GCT ACA AGC ACT GGA GAA GGT AGT AGA GAT Asp Asp Leu Ala Lys Val Ala Thr Ser Thr Gly Glu Gly Ser Arg Asp | 1333 |
| 40 | 130 135 140 | |
| 45 | ATT AAA ATC AGA CTA TTA GCA GGA TTA CTT AAG GAT GCT TCT OCA CTT Ile Lys Ile Arg Leu Leu Ala Gly Leu Leu Lys Asp Ala Ser Pro Leu | 1381 |
| | 145 150 155 | |
| 50 | GAG GCT AAA TAT TTA GTT AGA TTT GTT GAT GGT AGG TTA AGG GTT GGA Glu Ala Lys Tyr Leu Val Arg Phe Val Asp Gly Arg Leu Arg Val Gly | 1429 |
| | 160 165 170 | |

55

| | | |
|----|--|------|
| | ATA GGA GAC GCT ACA ATA TTA GAC GCA CTA GCA ATA ACT TTC GGC GGC | 1477 |
| | Ile Gly Asp Ala Thr Ile Leu Asp Ala Leu Ala Ile Thr Phe Gly Gly | |
| 5 | 175 180 185 190 | |
| | GGG CAG AAC TTT AGG CCT ATA GTT GAA AGA GCA TAC AAT TTG AGG GGC | 1525 |
| 10 | Gly Gln Asn Phe Arg Pro Ile Val Glu Arg Ala Tyr Asn Leu Arg Ala | |
| | 195 200 205 | |
| | GAT TTA GGT AAT ATT GCA AAA ATA TTA GCA AAT GGC GGA ATT GAA CAA | 1573 |
| 15 | Asp Leu Gly Asn Ile Ala Lys Ile Leu Ala Asn Gly Gly Ile Glu Gln | |
| | 210 215 220 | |
| | CTG AAG AAC ATT AAA OCT CAG CCA GGG ATT CCA ATA AGG CCA ATG CTA | 1621 |
| 20 | Leu Lys Asn Ile Lys Pro Gln Pro Gly Ile Pro Ile Arg Pro Met Leu | |
| | 225 230 235 | |
| | GCT GAA AGA TTA TCA GAC OCT GCA GAG ATG TTA AGT AAG GTA GGC AAT | 1669 |
| 25 | Ala Glu Arg Leu Ser Asp Pro Ala Glu Met Leu Ser Lys Val Gly Asn | |
| | 240 245 250 | |
| | ATT GCA TTG GTC GAT TAT AAG TAT GAT GGA GAG AGA GGA CAA ATT CAT | 1717 |
| 30 | Ile Ala Leu Val Asp Tyr Lys Tyr Asp Gly Glu Arg Gly Gln Ile His | |
| | 255 260 265 270 | |
| 35 | | |
| | AAA GCA GGA GAT AAA ATT TTC ATA TTC TCT AGA AGA TTG GAA AAT ATA | 1765 |
| 40 | Lys Ala Gly Asp Lys Ile Phe Ile Phe Ser Arg Arg Leu Glu Asn Ile | |
| | 275 280 285 | |
| | ACA AAT CAA TAT OCT GAC GTA GCA GAG TAT ATT TCA AAA TAC GTT AAA | 1813 |
| 45 | Thr Asn Gln Tyr Pro Asp Val Ala Glu Tyr Ile Ser Lys Tyr Val Lys | |
| | 290 295 300 | |
| 50 | | |
| 55 | | |

| | | |
|----|--|------|
| 5 | GGT AAT GAG TTC ATA GTT GAG GGA GAA ATA ATA OCA GTA GAT CCA GAA Gly Asn Glu Phe Ile Val Glu Gly Glu Ile Ile Pro Val Asp Pro Glu | 1861 |
| | 305 310 315 | |
| 10 | ACT GGT GAA ATG AGA CCT TTC CAA GAA TTG ATG CAC AGG AAA AGG AAG Thr Gly Glu Met Arg Pro Phe Gln Glu Leu Met His Arg Lys Arg Lys | 1909 |
| | 320 325 330 | |
| 15 | TCC GAT ATT CAC GAG GCT ATA AAA GAG TAT CCG GTT AAC GTA TTC TTA Ser Asp Ile His Glu Ala Ile Lys Glu Tyr Pro Val Asn Val Phe Leu | 1957 |
| | 335 340 345 350 | |
| 20 | TTT GAT TTA ATG TAC TAT GAA GGA GAA GAT TAT ACA GTA AAG CCG TTA Phe Asp Leu Met Tyr Tyr Glu Gly Glu Asp Tyr Thr Val Lys Pro Leu | 2005 |
| | 355 360 365 | |
| 25 | TCA GAA AGA CGT AAG AAG TTA GAA TCA ATA GTT GAA GAT AAC GAT TAT Ser Glu Arg Arg Lys Lys Leu Glu Ser Ile Val Glu Asp Asn Asp Tyr | 2053 |
| | 370 375 380 | |
| 30 | GTA CAT ATA GCC ACT CAT ATT ATA ACT GAT AAT GTA GAA AAA CTA AAG Val His Ile Ala Thr His Ile Ile Thr Asp Asn Val Glu Lys Leu Lys | 2101 |
| | 385 390 395 | |
| 35 | GAA TTC TTC TAT CAA GCA ATC TCT GAA GGC GCA GAA GGC GTA ATG GTA Glu Phe Phe Tyr Gln Ala Ile Ser Glu Gly Ala Glu Gly Val Met Val | 2149 |
| | 400 405 410 | |
| 40 | AAA TCT CTG GCA OCA GAT GCA ATA TAT CAA GCA GGA TCA AGA GGT TGG Lys Ser Leu Ala Pro Asp Ala Ile Tyr Gln Ala Gly Ser Arg Gly Trp | 2197 |
| 45 | 415 420 425 430 | |

50

55

| | | |
|----|---|------|
| | CTT TGG ATA AAG TTT AAG AGA GAT TAC CAG AGC GAA ATG GCA GAT ACT | 2245 |
| | Leu Trp Ile Lys Phe Lys Arg Asp Tyr Gln Ser Glu Met Ala Asp Thr | |
| 5 | 435 440 445 | |
| | GTA GAC CTG GTA ATG GTA GGT GCT TTT CAC GGT AAA GGA AGG AAA GGA | 2293 |
| | Val Asp Leu Val Met Val Gly Ala Phe His Gly Lys Gly Arg Lys Gly | |
| 10 | 450 455 460 | |
| | GGC AAA TAT AGT TCC TTC TTA ATG GCT GCA TAT AAT OCA GAC AAG GAC | 2341 |
| | Gly Lys Tyr Ser Ser Phe Leu Met Ala Ala Tyr Asn Pro Asp Lys Asp | |
| 15 | 465 470 475 | |
| | GTT TTC GAG ACT GTA TGT AAA GTT GCT TCT GGT TTT ACT GAT GCA GAA | 2389 |
| | Val Phe Glu Thr Val Cys Lys Val Ala Ser Gly Phe Thr Asp Ala Glu | |
| 20 | 480 485 490 | |
| | CTT GAC GAT TTA CAG AAG AAA ATT GGC GAG TTA AAA AGA GAC ACA OCT | 2437 |
| | Leu Asp Asp Leu Gln Lys Lys Ile Ala Glu Leu Lys Arg Asp Thr Pro | |
| 25 | 495 500 505 510 | |
| | CAT OCA AGA GTT GTA TCA ACC ATG GTT OCA GAT GTT TGG TTA ACA OCA | 2485 |
| | His Pro Arg Val Val Ser Thr Met Val Pro Asp Val Trp Leu Thr Pro | |
| 30 | 515 520 525 | |
| | GCA CTT GTT GGC GAA ATA ATT GGA GCA GAA ATA ACT ATA TCT OCT CTA | 2533 |
| | Ala Leu Val Ala Glu Ile Ile Gly Ala Glu Ile Thr Ile Ser Pro Leu | |
| 35 | 530 535 540 | |
| | CAT ACT TGT TGC AAA GAT CAG TAT GCA GAA GGT GGT TTA TCA ATA AGA | 2581 |
| | His Thr Cys Cys Lys Asp Gln Tyr Ala Glu Gly Gly Leu Ser Ile Arg | |
| 40 | 545 550 555 | |
| 45 | | |
| 50 | | |
| 55 | | |

50

(2) INFORMATION ZU SEQ ID NO: 6:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 600 Aminosäuren

(B) ART: Aminosäure

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

Met Glu Phe Lys Ile Ile Ala Glu Tyr Phe Asp Arg Leu Glu Lys Ile

1 5 10 15

Ser Ser Arg Leu Gln Leu Thr Ser Leu Leu Ala Asp Leu Phe Lys Lys

20 25 30

Thr Asp Lys Asn Val Ile Asp Lys Val Val Tyr Ile Ile Gln Gly Lys

35 40 45

Leu Trp Pro Asp Phe Leu Gly Met Pro Glu Leu Gly Ile Gly Glu Lys

50 55 60

Phe Leu Ile Arg Ala Leu Ser Ile Ala Thr Ser Val Ser Asp Asp Glu

65 70 75 80

Ile Glu Lys Met Tyr Lys Ser Val Gly Asp Leu Gly Gln Val Ala Phe

85 90 95

Asp Ile Lys Gln Lys Gln Gln Ser Ala Ser Ile Leu Ala Phe Leu Gly

100 105 110

Ala Gln Lys Ala Ser Lys Pro Leu Thr Val Glu Lys Val Tyr Asp Asp
115 120 125

Leu Ala Lys Val Ala Thr Ser Thr Gly Glu Gly Ser Arg Asp Ile Lys
130 135 140

Ile Arg Leu Leu Ala Gly Leu Leu Lys Asp Ala Ser Pro Leu Glu Ala
145 150 155 160

Lys Tyr Leu Val Arg Phe Val Asp Gly Arg Leu Arg Val Gly Ile Gly
165 170 175

Asp Ala Thr Ile Leu Asp Ala Leu Ala Ile Thr Phe Gly Gly Gly Gln
180 185 190

Asn Phe Arg Pro Ile Val Glu Arg Ala Tyr Asn Leu Arg Ala Asp Leu
195 200 205

Gly Asn Ile Ala Lys Ile Leu Ala Asn Gly Gly Ile Glu Gln Leu Lys
210 215 220

Asn Ile Lys Pro Gln Pro Gly Ile Pro Ile Arg Pro Met Leu Ala Glu
225 230 235 240

Arg Leu Ser Asp Pro Ala Glu Met Leu Ser Lys Val Gly Asn Ile Ala
245 250 255

Leu Val Asp Tyr Lys Tyr Asp Gly Glu Arg Gly Gln Ile His Lys Ala
260 265 270

Gly Asp Lys Ile Phe Ile Phe Ser Arg Arg Leu Glu Asn Ile Thr Asn
275 280 285

Gln Tyr Pro Asp Val Ala Glu Tyr Ile Ser Lys Tyr Val Lys Gly Asn
 290 295 300

Glu Phe Ile Val Glu Gly Glu Ile Ile Pro Val Asp Pro Glu Thr Gly
 305 310 315 320

Glu Met Arg Pro Phe Gln Glu Leu Met His Arg Lys Arg Lys Ser Asp
 325 330 335

Ile His Glu Ala Ile Lys Glu Tyr Pro Val Asn Val Phe Leu Phe Asp
 340 345 350

Leu Met Tyr Tyr Glu Gly Glu Asp Tyr Thr Val Lys Pro Leu Ser Glu
 355 360 365

Arg Arg Lys Lys Leu Glu Ser Ile Val Glu Asp Asn Asp Tyr Val His
 370 375 380

Ile Ala Thr His Ile Ile Thr Asp Asn Val Glu Lys Leu Lys Glu Phe
 385 390 395 400

Phe Tyr Gln Ala Ile Ser Glu Gly Ala Glu Gly Val Met Val Lys Ser
 405 410 415

Leu Ala Pro Asp Ala Ile Tyr Gln Ala Gly Ser Arg Gly Trp Leu Trp
 420 425 430

Ile Lys Phe Lys Arg Asp Tyr Gln Ser Glu Met Ala Asp Thr Val Asp
 435 440 445

Leu Val Met Val Gly Ala Phe His Gly Lys Gly Arg Lys Gly Gly Lys
 450 455 460

Tyr Ser Ser Phe Leu Met Ala Ala Tyr Asn Pro Asp Lys Asp Val Phe
 465 470 475 480

Glu Thr Val Cys Lys Val Ala Ser Gly Phe Thr Asp Ala Glu Leu Asp
 485 490 495

Asp Leu Gln Lys Lys Ile Ala Glu Leu Lys Arg Asp Thr Pro His Pro
 500 505 510

Arg Val Val Ser Thr Met Val Pro Asp Val Trp Leu Thr Pro Ala Leu
 515 520 525

Val Ala Glu Ile Ile Gly Ala Glu Ile Thr Ile Ser Pro Leu His Thr
 530 535 540

Cys Cys Lys Asp Gln Tyr Ala Glu Gly Gly Leu Ser Ile Arg Phe Pro
 545 550 555 560

Arg Phe Ile Arg Trp Arg Pro Asp Lys Ser Pro Glu Asp Ala Thr Thr
 565 570 575

Asn Arg Glu Ile Leu Glu Met Tyr Lys Ser Gln Leu Lys Lys Ile Glu
 580 585 590

Glu Lys Pro Ser Asp Gln Ser Val
 595 600

(2) INFORMATION ZU SEQ ID NO: 7:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 17 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

ACTTTGGCCA TTGGAAT

17

(2) INFORMATION ZU SEQ ID NO: 8:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 17 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

CTCTTAAATC GACCCCG

17

(2) INFORMATION ZU SEQ ID NO: 9:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 38 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

GCGGATCCAT CGAAGGTAGA ATGGAGTTTA AAATTATA

38

(2) INFORMATION ZU SEQ ID NO: 10:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 30 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

TATGGATCCT TATACACTCT GATCTGAOCC

30

(2) INFORMATION ZU SEQ ID NO: 11:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 17 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

GTITTTCCAG TCAAGAC

17

(2) INFORMATION ZU SEQ ID NO: 12:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 18 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

CACACAGGAA ACAGCTAT

18

(2) INFORMATION ZU SEQ ID NO: 13:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 17 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

ATTAACCCCTC ACTAAAG

17

(2) INFORMATION ZU SEQ ID NO: 14:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 17 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

AATACGACTC ACTATAG

17

Patentansprüche

1. Thermostabile Ligase erhältlich aus Bakterien der Gruppe der Archaeobakterien.

2. Ligase gemäß Anspruch 1, erhältlich aus Bakterien der Ordnung der Sulfolobales.

3. Ligase enthaltend

- die Aminosäuresequenz der FIG 1 oder
- einen Teil der Aminosäuresequenz der FIG 1 mit mehr als 50 Aminosäuren Länge, der Ligaseaktivität aufweist, oder
- eine Aminosäuresequenz, die mindestens zu 40 % homolog zu einer der genannten Sequenzen ist und Ligaseaktivität aufweist.

4. Ligase gemäß Anspruch 1, dadurch gekennzeichnet, daß sie ATP-abhängig ist.

5. Ligase gemäß Anspruch 1, dadurch gekennzeichnet, daß sie weniger als 36 % Homologie zu der Aminosäuresequenz von eukaryonten DNA-Ligasen aufweist.

6. Ligase gemäß Anspruch 3, mit einem Molekulargewicht von ca. 67.6 kD.
7. Ligase gemäß einem der vorangehenden Ansprüche, dadurch gekennzeichnet, daß sie rekombinant hergestellt ist.
- 5 8. Ligase gemäß Anspruch 2, erhältlich aus Desulfurolobus ambivalens.
9. Thermostabile DNA-Ligase, erhältlich aus dem Expressionstamm mit der Hinterlegungsnummer DSM 7070.
- 10 10. Nukleinsäuresequenz, kodierend für eine thermostabile Ligase aus Archaeobakterien.
11. Nukleinsäuresequenz nach Anspruch 10, enthaltend
- die Nukleotidsequenz der FIG 1 oder
 - 15 - einen Teil davon, der für eine Ligase kodiert und mindestens 150 bp lang ist, oder
 - eine Sequenz, die mindestens zu 40 % homolog zu einer der genannten Sequenzen ist und für eine Ligase kodiert.
12. Nukleinsäuresequenz, deren Komplement unter stringenten Bedingungen mit der Nukleinsäure nach
- 20 FIG 1 hybridisiert.
13. Nukleinsäuresequenz nach Anspruch 12, deren Komplement unter stringenten Bedingungen mit dem 3382 bp langen XbaI/Hind III-Fragment aus Desulfurolobus ambivalens hybridisiert.
- 25 14. Nukleotidsequenz von mindestens 150 Nukleotiden Länge, enthalten in oder komplementär zu dem 1800 bp langen BamHI-Fragment aus DSM 7070.
15. Verfahren zum Nachweis einer Nukleinsäure durch Hybridisierung zweier Oligonukleotide mit benach-
- 30 barten Sequenzen der Nukleinsäure, Ligation und Nachweis des Ligationsproduktes, dadurch gekennzeichnet, daß die Ligation von der Ligase gemäß Anspruch 1 mit ATP als Cofaktor vorgenommen wird.
16. Expressionsclon M15-pQE-L6 mit der Hinterlegungsnummer DSM 7070.

35

40

45

50

55

FIG. 1a

```

1 - AAGCTTAAGAATCCTGACTTATTCTTTGAAAAAATTGAAAAGCAAGGTGGGAAAGTTCTA - 60
  K L K N P D L F F E K I E K Q G G K V L
61 - TATACTGATAAGAAAGATAGCCGTAGATGTCGATTCTTTAATAATTTTGTAAAGAAA - 120
  Y T D K E R I A V D V D F F N N F V K K
121 - CTTTCAGATATTCACACTTCAGATGATGTAGAAGCGCAAAATACCTCACTAAGCAAGAA - 180
  L S D I H T S D D V E A Q K Y L T K Q E
181 - CATAAATTATCCAAAAATTGAGACAATTGGGAATAATATACTTCGATAATAATACTAAG - 240
  H K L F Q K L R Q L G I I Y F D N N T K
241 - GCATGGAAGTTATCGGAATTTAAATTTTGGTAAAGTTACGCCTTGTGTCTTGGTATTT - 300
  A W K L S E F K F W *
  * R F K F K P L T V G Q Q G Q Y K
301 - GCCATGATAAGGATTCTCTACTGGTGTAAATGTTCTGGCGAAAATGAGGTGAAGAAAGCG - 360
  G H Y P N E V P T L T R A F I L H L F R
361 - AGTCCCCCTCTTTATCTTCACGGGAAAAGCAGAACCTAGTACTTCTATTGTGAGTTGTCC - 420
  T G R K I K V P F A S G L V E I T L Q G
421 - TTCAAAGCCTGCATCTACGATAGTCGGTGGTACAAAAGACCCAGCCTGGCAAAAGAAGA - 480
  E F G A D V I T P P V F L G L R A F S S
481 - CCTTAAGTTGACGAAAGCCATTACGTCGTTCCGGCAGTTTAAACATACTCCTCAGTAACTAA - 540
  R L N V F A M V D N P L K V Y E E T V L
541 - CAGTACATGTTGTTAGGTAATGATGAATTCATCTCCTTTCTCTATCTCATAAAATGA - 600
  L V H E N P Y I I F E D G K E I E Y F S
601 - ACGCGGGTCTTTTCCATCTTCATATATTTTCGTCAGTCTTCTTAAAGCGGGCTATCTCTCC - 660
  R P D K G D E Y I E D T K K F R A I E G
661 - TCCTACTCTTAAATCGACCCCGTCTCTCTTATAGTATCTTGAGTAAGAGGAGAGATTAC - 720
  G V R L D V G N E R I T D Q T L P S I V
721 - TATCCAACCCTTTTCTAGGTAGTATTTTAAATCTCTATCTCCAAGAATCATAAATATCTA - 780
  I W G K E L Y Y K L D R D G L I M ← pDam-L3-ORF3
781 - CTATTTTTCATCTTGCAGAATATATTTTATAAGGAAAATGGCTTCAGAGTAACTCTGAT - 840
841 - ACTTCATCCTTCTGGCTTGCCGGTTTTTCATCATCAAAATATATTATATCGTAATTCAACT - 900
901 - TTTATTTATGGAGTTTAAATTTATAGCTGAATCTTTGATAGACTAGAAAAGATCTCTTC - 960
  lig → M E F K I I A E Y F D R L E K I S S
961 - TAGGCTTCAACTTACATCATTACTTGCAGATCTTTTAAAGAAAATGACAAAAATGTTAT - 1020
  R L Q L T S L L A D L F K K T D K N V I
1021 - TGATAAAGTTGTATATATAATTCAGGGTAAACTATGGCCAGACTTCTTAGGAATGCCAGA - 1080
  D K V V Y I I Q G K L W P D F L G M P E
1081 - ATTAGGAATAGGTGAAAAATTCCTTATAAGAGCATTATCAATAGCTACAAGCGTAAGCGA - 1140
  L G I G E K F L I R A L S I A T S V S D
1141 - TGATGAAATAGAAAAATGTATAAATCAGTAGGAGATTGTTGGGCAAGTAGCTTTTGATAT - 1200
  D E I E K M Y K S V G D L G Q V A F D I
1201 - AAAGCAAAACAGCAGTCTGCCAGTATTTTAGCGTTTTTAGGAGCGCAAAAGCTTCAAA - 1260
  K Q K Q Q S A S I L A F L G A Q K A S K
1261 - GCCGCTAACTGTAGAAAAGTTTATGATGATTAGCAAAAGTTGCTACAAGCACTGGAGA - 1320
  P L T V E K V Y D D L A K V A T S T G E
1321 - AGGTAGTAGATATTAAATCAGACTATTAGCAGGATTACTTAAGGATGCTTCTCCACT - 1380
  G S R D I K I R L L A G L L K D A S P L
1381 - TGAGGCTAAATATTAGTTAGATTGTTGATGGTAGGTTAAGGGTTGGAATAGGAGACGC - 1440
  E A K Y L V R F V D G R L R V G I G D A
1441 - TACAATATTAGACGCACTAGCAATAACTTTCCGGCGGGCAGAACTTTAGGCCTATAGT - 1500
  T I L D A L A I T F G G G Q N F R P I V
1501 - TGAAAGAGCATACAATTTGAGGGCCGATTTAGGTAATATTGCAAAAATATTAGCAATGG - 1560
  E R A Y N L R A D L G N I A K I L A N G
1561 - CGGAATTGAACAACTGAAGAACATTAACCTCAGCCAGGAGTTCCAATAGGCCAATGCT - 1620
  G I E Q L K N I K P Q P G I P I R P M L
1621 - AGCTGAAAGATTATCAGACCCTGCAGAGATGTTAAGTAAGGTAGGCAATATTGCATTGGT - 1680
  A E R L S D P A E M L S K V G N I A L V

```

FIG. 1b

1681 - CGATTATAAGTATGATGGAGAGAGAGGACAAATTCATAAAGCAGGAGATAAAATTTTCAT - 1740
 D Y K Y D G E R G Q I H K A G D K I F I
 1741 - ATTCTCTAGAAGATTGGAAAATATAACAAATCAATATCCTGACGTAGCAGAGTATATTTC - 1800
 F S R R L E N I T N Q Y P D V A E Y I S
 1801 - AAAATACGTTAAAGGTAATGAGTTCATAGTTGAGGGAGAAATAATACCAGTAGATCCAGA - 1860
 K Y V K G N E F I V E G E I I P V D P E
 1861 - AACTGGTGAAATGAGACCTTTCCAAGAATTGATGCACAGGAAAAGGAAGTCCGATATTCA - 1920
 T G E M R P F Q E L M H R K R K S D I H
 1921 - CGAGGCTATAAAAGAGTATCCGGTTAACGTATTCTTATTGATTAAATGTACTATGAAGG - 1980
 E A I K E Y P V N V F L F D L M Y Y E G
 1981 - AGAAGATTATACAGTAAAGCCGTTATCAGAAAGCGTAAGAAGTTAGAATCAATAGTTGA - 2040
 E D Y T V K P L S E R R K K L E S I V E
 2041 - AGATAACGATTATGTACATATAGCCACTCATATTATAACTGATAATGTAGAAAACTAAA - 2100
 D N D Y V H I A T H I I T D N V E K L K
 2101 - GGAATTCTTCTATCAAGCAATCTCTGAAGGCGCAGAAAGCGTAATGGTAAAATCTCTGGC - 2160
 E F F Y Q A I S E G A E G V M V K S L A
 2161 - ACCAGATGCAATATATCAAGCAGGATCAAGAGGTTGGCTTTGGATAAAGTTTAAAGAGAGA - 2220
 P D A I Y Q A G S R G W L W I K F K R D
 2221 - TTACCAGAGCGAAATGGCAGATACTGTAGACCTGGTAATGGTAGGTGCTTTTCACGGTAA - 2280
 Y Q S E M A D T V D L V M V G A F H G K
 2281 - AGGAAGGAAAGGAGGCAAAATATAGTTCCTTCTTAATGGCTGCATATAATCCAGACAAGGA - 2340
 G R K G G K Y S S F L M A A Y N P D K D
 2341 - CGTTTTCGAGACTGTATGTAAAGTTGCTTCTGCTTTTACTGATGCAGAACTTGACGATT - 2400
 V F E T V C K V A S G F T D A E L D D L
 2401 - ACAGAAGAAAATTGCCGAGTTAAAAAGAGACACCTCATCCAAGAGTTGTATCAACCAT - 2460
 Q K K I A E L K R D T P H P R V V S T M
 2461 - GGTTCCAGATGTTTGGTTAACACCAGCACTTGTGGCGAAATAATTGGAGCAGAAATAAC - 2520
 V P D V W L T P A L V A E I I G A E I T
 2521 - TATATCTCCTCTACATACTTGTGCAAGATCAGTATGCAGAAGGTGGTTTATCAATAAG - 2580
 I S P L H T C C K D Q Y A E G G L S I R
 2581 - ATTCCCAAGATTTATAAGGTGGAGACCAGATAAGAGTCCAGAAGATGCAACTACTAATCG - 2640
 F P R F I R W R P D K S P E D A T T N R
 2641 - TGAAATCTTGAGATGTATAAGAGCCAATTAAAGAAGATTGAAGAAAAGCCGTCAGATCA - 2700
 E I L E M Y K S Q L K K I E E K P S D Q
 2701 - GAGTGATAAGTTTTTATAAAATAACTCATAGTTATGTTTTGTGTTCTTCGATATTAAGA - 2760
 S V *
 2761 - AGTCTGGTGCAACTTTTTAGGTAACAACCTTTACAGTAGATGGACATTATGAGCGAAATT - 2820
 T TAGGGTTGTAACATTTTTTCATGCTGACCATATTTTGCAATTAAATAAGAGCGTATCAT - 2880
 2881 - CTTGTATTGGGATTATATCAACTCCTCCAACCTCTGAAGTTCTCTCAGTTTATAGGGTATG - 2940
 2941 - ATATACCTAAGAAAAAGCAATGGGAATAAGTTATGGTGTAATAAATTGCAATAGAAGATG - 3000
 pDam-L3-ORF2 → M G I S Y G V K I A I E D E
 3001 - AGAAAATAGAAGTATGTTACAGGAGATCATGTATTTGGAGCCGCTCAAGTAGTTGTAACAA - 3060
 K I E L V Q A D H V F G A A Q V V V T N
 3061 - ATTCCAATGGCGAAAGTGATAGTTATAGTGGAGATTTTAAAGATCCGGGTAAAGGTACTC - 3120
 S N G E S V G Y T G D F K N P G K G T P
 3121 - CTATACTTAATACTGACGTATTAATTGTTGATACTACCTACGGAAAGCCGACATTTAGAA - 3180
 I L N T D V L I V D T T Y G K P T F R R
 3181 - GAAAATTTAGAGATGAAGTTGAAGTCTTATTTTCAGATTACGTAAACGATTTCATTAATT - 3240
 K F R D E V E V L F S D Y V N D S L I Y
 3241 - ATGGACCAGTAAGAGTTTATGCTTACTATGGAAGATTCAAGAGGCAATGAAAATACTGA - 3300
 G P V R V Y A Y Y G K I Q E A M K I L R
 3301 - GGAAAATGGAATTTCCGGCCCTTTTATAGTAGATGGTAAAATAAAGAAGTTACAGATA - 3360
 K N G I S A P F I V D G K I K E V T D I
 3361 - TTGCAATAAAGTATGGTCTAGA
 A I K Y G L →

FIG. 2

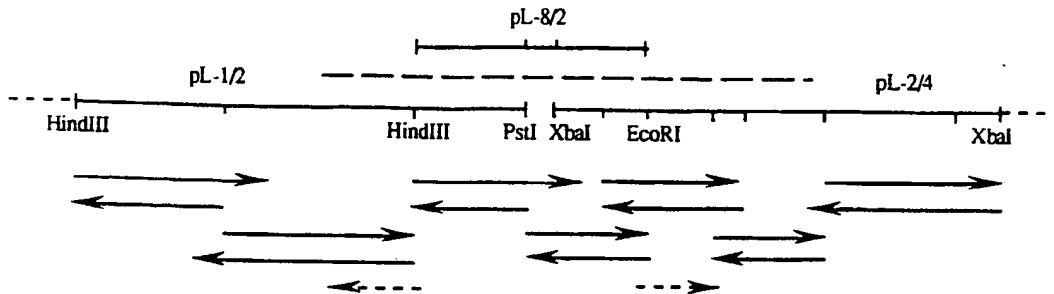


FIG. 3

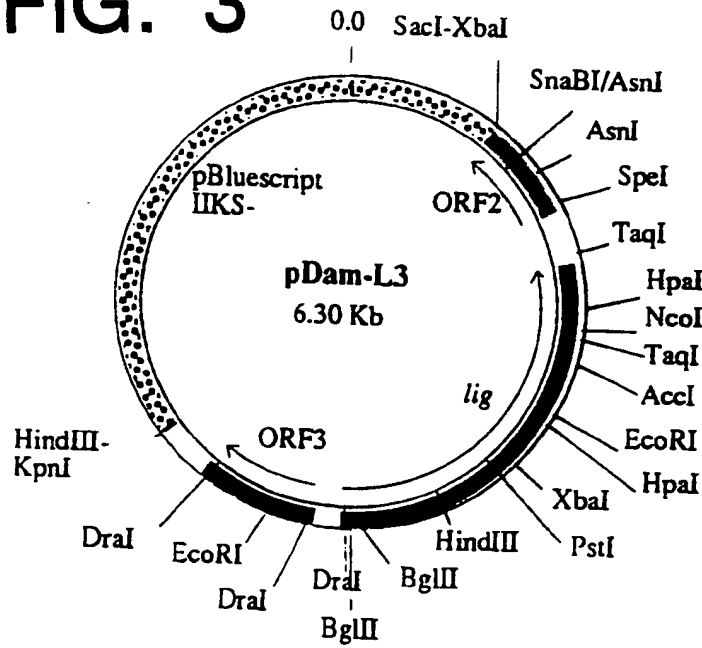


FIG. 4a

| | | | | | | | | | |
|---------|-----|------------|------------|-------------|-------------|-------------|-------------|------------|-------------------------------|
| H. sap. | 1 | MORSIMSFFH | PKKEGKAKKP | EKEASNSSRE | TEPPPKAALK | EWNGVVSESD | SPVKRPGRKA | 60 | - 60 |
| H. sap. | 61 | ARVLGSEGE | EDEALSPAKG | QKPALDCSQV | SPPRPATSPE | NNASL SDTS | P.MDSSPSGI | 120 | - 118 |
| S. cer. | | MRRLTGC | LLSSARPL | KSRLPLIMS | SSLPSSAGKK | | | 35 | - 35 |
| S. pom. | | MRTVFSQI | PRFKQVNYI | RMSTRQSDIS | NFFISSASHK | | | 38 | - 38 |
| Kons. | | mr | sq | p | s | l | sd | S | SSas k |
| H. sap. | 121 | PKRRTARKOL | PKRTIQEVLE | EQSEDEDEA | KRKKEEEEE | TPKESLTA | VATEKEGEDG | 180 | - 178 |
| S. cer. | | PKOATL | | | | AR | FFTSMKNKPT | 53 | - 53 |
| S. pom. | | SEHVEV | | | | SQ | SSSDSKNVDG | 56 | - 56 |
| Kons. | | pk | t | | | a | t | kn | dg |
| H. sap. | 181 | DQPTTPPKPL | KTSKAETPTE | SVS.EPEVAT | KQELQEEEEE | TKPPRRAPKT | LSSFFTTPRKP | 240 | - 237 |
| S. cer. | | EGTPSPKKSS | KHML.EERMD | NVSGEEEYAT | KKLKQTAVTH | TVAAPSSMG | NFS..SIPSS | 110 | - 110 |
| S. pom. | | RSTSEKRKVE | SVKLVDESH | N..NHDDTGT | QNVRE..NN | IVSEAKKQKT | LGS..SSSSS | 110 | - 110 |
| Kons. | | t | p | k | k | l | ee | nvs | e e at k q tv t l s s ss |
| B. tau. | 241 | | | | | | V | PTLAVAR | 300 |
| D. amb. | | | | | | | M | EFKIIAEYFD | 11 |
| H. sap. | | AVKKEVKEEE | PGAPGKEGAA | EGPLDPSPGYN | PAKNNYHPVE | DACWKPGQKV | PYLAVARTFE | 297 | - 297 |
| S. cer. | | APSSGVADSP | QOSQRLVGEV | EDALSS..N | N.NDHYSSN | | PYSEVCEVEN | 156 | - 156 |
| S. pom. | | DAVSSNNDSG | ASTIPLPIK | EPPLS..N | ARNDKLKGH | | TFAEMVKAFT | 157 | - 157 |
| Vacc. | | | | | | | MT | SLREFRKLCC | 12 |
| Kons. | | a | s | vds | p | g | E | pl | s N nd y e f |
| D. amb. | 301 | RLEKISSRLQ | LTSLLADLFK | KTDKNVIDKV | V...YIIQK | LWPDFLGMPE | LGIGEFILIR | 360 | - 68 |
| H. sap. | | KIEEVSARLR | MVETLSNLLR | SVVALSPDDL | LPVLYLSLNH | LGPP..QOGL | LGVDGVLK | 356 | - 356 |
| S. cer. | | KIEAISRLE | IIRICSDFFI | KIMKQSSKNV | IPITYLFINR | LGPDYEAGLE | LGLGENLLMK | 216 | - 216 |
| S. pom. | | KIENTSKRLE | IIDIMGTFFF | GILRDHPSDL | LACVYLSINK | LGPDY..SGLE | LGIGESIIMK | 216 | - 216 |
| Vacc. | | KIYHASYKE | KSKLIRDFIT | DRDDK.... | YLIKL | LLPGLDDRI | YNMNDKQIIR | 62 | - 62 |
| Kons. | | kie | S | rle | YL | in | LgPd | le | lg ge i k |
| D. amb. | 361 | ALSIATSVSD | DEIERMYKSV | GDLGOVAFDI | KQKQOSASIL | AFLGAQKASK | PLTVEKVYDD | 420 | - 128 |
| H. sap. | | AVAQATGRQL | ESVRAEAAEK | GDVGLVAENS | RSTORLMLP | PP | PLTASGVFSK | 407 | - 407 |
| S. cer. | | TISETCGKSM | SOIKLKYKDI | GDLGEIAMGA | RNVQPTMFK | PK | PLTVGEVEKN | 267 | - 267 |
| S. pom. | | AIGESTGOTL | QOIKLSFHKV | GDLGLVAQTS | RQNOPTMFK | PA | ALTIPFLFDS | 267 | - 267 |
| Vacc. | | LYSIIFKQSQ | EDM.....L | QDLGYGYIG | .DTIRTEFFE | NTEIRPRDKS | ILTLEDVDSF | 114 | - 114 |
| T4 | | | | | | | M | ILKILNEIAS | 11 |
| Kons. | | | i | | gdIG | va | q | mfk | p pLc vf |
| D. amb. | 421 | LAKVATSTGE | GSRDIKIRLL | AGLLKDASPL | EAKYLVRVFD | GRLRVGIGDA | TILDALAITF | 480 | - 188 |
| H. sap. | | FRDIARLTGS | ASTAKRIDII | KGLFVACRHS | EARFIARSLS | GRLRLGLAEQ | SVLAALSQAV | 467 | - 467 |
| S. cer. | | LRAIAKTQK | DSQLKKMKLI | KRMLTACKGI | EAKFLIRSLE | SKLRIGLAEK | TVLISLSKAL | 327 | - 327 |
| S. pom. | | LKKIAQMSGN | QSQNRKIGVI | KRLSSCEGA | EPKYLRALAE | GKLRLQLAEK | TILVALANAT | 327 | - 327 |
| Vacc. | | LTTLSSTVKE | SHQ...IKLL | TDIASVCTCN | DLKCVMLID | KOLKIKAGPR | YVLNAISPNA | 171 | - 171 |
| T4 | | IGSTKQKQAI | LEKNKDNELL | KRVYRLTYSR | GLQYYIKKWP | .KPGIATQSF | GMLTLTDMLD | 70 | - 70 |
| Kons. | | l | ia | tq | sq | kki | l* | krll | c eakyliir klriglae cvL als a |
| D. amb. | 481 | GGGQNF | P | | | | | | 540 |
| H. sap. | | SLTPPGQEFF | PAMVDAGKKG | TAEARKTWLE | EQGMILKQTF | CEVPDLDRII | PVLEHG | 524 | - 524 |
| S. cer. | | LLHDENRE | DSP | DKDVPMOVLE | SAQOKIRDAF | CQVNYEIVI | NSCLEHG | 375 | - 375 |
| S. pom. | | AQYHADKN | G | EKLSQODRIE | GEQ..ILRDVY | COLPSYDLIV | PHLIEHG | 372 | - 372 |
| Vacc. | | Y | | | | DVF | RKSNMLKEII | ENASKQNLDS | 195 |
| T4 | | FIEFTLAT | | RKLIGNAAIE | ELTGYITDG | .KKDDVEVLR | RVMMDLECG | 116 | - 116 |
| Kons. | | | | e | i | d | f | c | pd1 ii llehg |

FIG. 4b

FIG. 5

| Species | S. p. | S. c. | H. s. | D. a. | Vacc. | T4DNA | T6 | T7 | T3 | E. c. | T. th. | T4RNA | S. c. 1 |
|--|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| <i>Sch. pombe</i> (767) | 421,21 | 46,2 | 43,6 | 35,1 | 30,1 | 20,5 | 20,5 | 20,3 | 24,6 | 18,0 | 21,4 | 15,3 | 16,6 |
| <i>S. cerevisiae</i> (755) | 210,57 | 397,20 | 42,0 | 34,3 | 29,7 | 21,3 | 22,0 | 23,4 | 22,1 | 16,5 | 20,3 | 18,2 | 15,3 |
| <i>H. sapiens</i> (919) | 190,52 | 174,42 | 393,44 | 35,5 | 27,0 | 23,4 | 23,9 | 22,6 | 26,0 | 18,4 | 20,5 | 16,3 | 16,8 |
| <i>D. ambivalens</i> (600) | 86,82 | 84,01 | 79,28 | 270,56 | 32,8 | 20,7 | 20,6 | 23,7 | 20,2 | 20,4 | 16,8 | 16,9 | 16,0 |
| Vaccinia virus (552) | 52,94 | 53,34 | 51,22 | 48,54 | 266,08 | 22,8 | 23,7 | 21,7 | 22,8 | 15,3 | 18,0 | 17,4 | 14,4 |
| T4 DNA ligase (487) | 1,72 | 0,13 | -0,24 | -0,43 | 4,38 | 244,25 | 98,8 | 21,6 | 21,5 | 20,5 | 20,3 | 13,7 | 15,1 |
| T6 (487) | 1,52 | 0,11 | -0,08 | -0,49 | 4,82 | 231,92 | 239,88 | 21,3 | 21,8 | 20,7 | 20,0 | 13,1 | 15,3 |
| T7 (359) | 3,90 | 4,61 | 0,94 | 0,43 | 4,88 | 0,50 | 0,32 | 212,73 | 72,1 | 17,8 | 17,8 | 19,3 | 14,4 |
| T3 (346) | 1,33 | 0,16 | 1,34 | 0,48 | 8,32 | -1,41 | -1,42 | 121,92 | 195,81 | 20,8 | 20,1 | 20,0 | 22,3 |
| <i>E. coli</i> (672) | 1,05 | 2,36 | 2,12 | 0,19 | -0,98 | -0,36 | -0,46 | -1,16 | -0,32 | 334,90 | 46,7 | 18,2 | 22,7 |
| <i>Th. thermophilus</i> (678) | -0,61 | 0,81 | 0,64 | -0,88 | 0,58 | 0,01 | -0,04 | -1,44 | 1,90 | 141,84 | 268,24 | 16,2 | 15,5 |
| T4 RNA ligase (374) | 0,90 | 0,65 | -0,48 | -0,90 | 1,25 | 0,06 | 0,06 | 0,56 | 0,14 | 1,27 | 1,27 | 183,09 | 17,0 |
| <i>S. cerevisiae tRNA ligase</i> (827) | 0,96 | 1,86 | 0,54 | 0,10 | -0,76 | 0,46 | 0,71 | -0,86 | -1,02 | 0,95 | -0,18 | 1,62 | 428,25 |

ISDOCID: <EP_____0571880A1_I_>

↓

M E F K I

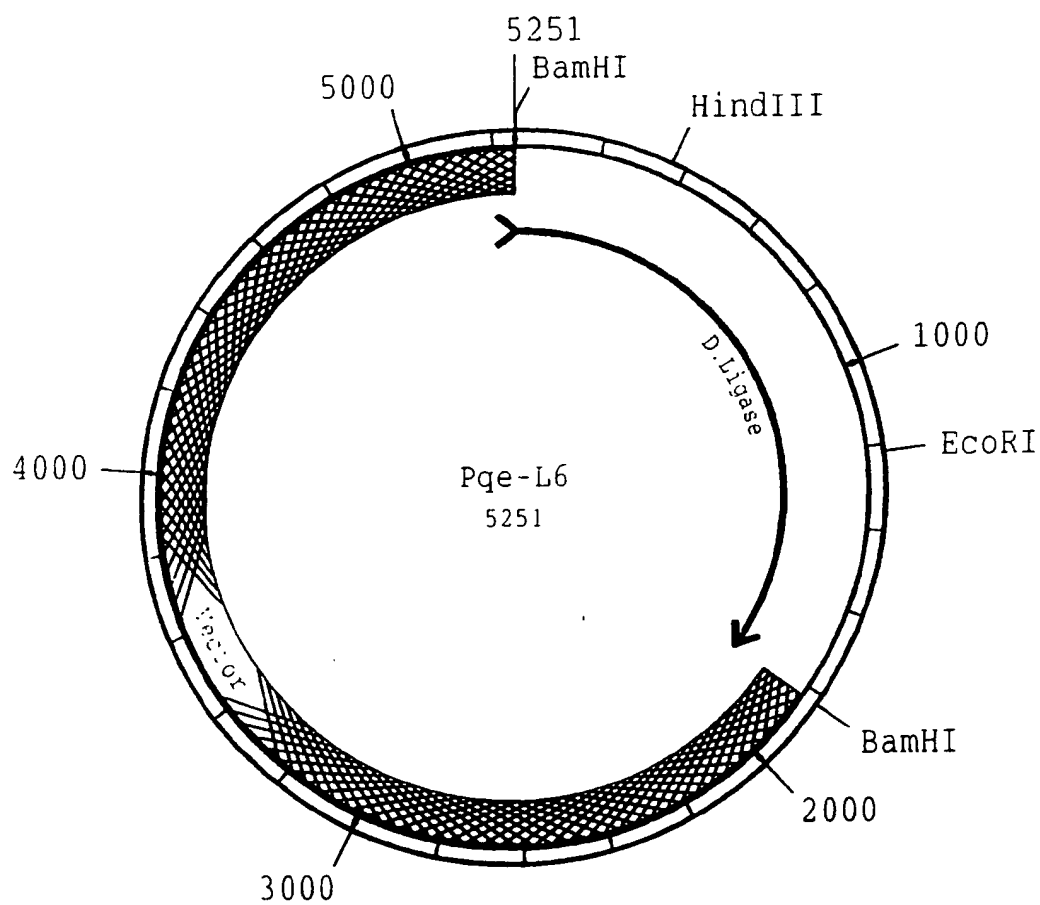
TTAT AAAATAACTC

CTTTT TACCTAACAA

ionssignal

CATT TTCATGCTGA

FIG. 7





Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung

EP 93 10 8104
PAGE1

| EINSCHLÄGIGE DOKUMENTE | | | |
|--|--|---|--|
| Kategorie | Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile | Betrifft Anspruch | KLASSIFIKATION DER ANMELDUNG (Int. Cl.5) |
| X,0 | BIOCHEMICAL SOCIETY SYMPOSIA Bd. 58, 1992, LONDON Seiten 149 - 169 DANSON, M. J., D. W. HOUGH AND G. G. LUNT (ED. 'THE ARCHAEBACTERIA: BIOCHEMISTRY AND BIOTECHNOLOGY' & SYMPOSIUM, EDINBURGH, SCOTLAND, UK, SEPTEMBER 1991 * das ganze Dokument * | 1,5,7 | C12N9/00 C12Q1/68 C12N15/52 |
| Y | --- | 10,15 | |
| D,Y | WO-A-9 117 239 (CORNELL RESEARCH FOUNDATION, INC.) * das ganze Dokument * | 10,15 | |
| A | --- | 1 | |
| A | WO-A-8 705 907 (INSTITUT PASTEUR) * Seite 4, Zeile 6 - Seite 7, Zeile 20 * | 10,15 | |
| A,D | FEMS MICROBIOLOGY REVIEWS Bd. 75, Nr. 2-3, 1990, AMSTERDAM NL Seiten 117 - 124 Stetter K O; Fiala G; Huber G; Huber R; Segerer A 'Hyperthermophilic microorganisms - hyper-thermophilic bacterium taxonomy; review' * das ganze Dokument * | 5 | |
| | | | RECHERCHIERTE SACHGEBIETE (Int. Cl.5) |
| | | | C12N C12Q |
| T | ADVANCES IN BIOCHEMICAL ENGINEERING/BIOTECHNOLOGY Bd. 45, 1992, BERLIN DE Seiten 57 - 98 COOLBEAR T; DANIEL RM; MORGAN HW 'The enzymes from extreme thermophiles: bacterial sources, thermostabilities and industrial relevance.' * das ganze Dokument * * Seite 87, Zeile 10 - Zeile 37 * --- -/-- | 1 | |
| Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt | | | |
| Recherchenort BERLIN | | Abschlußdatum der Recherche 11 AUGUST 1993 | Prüfer GURDJIAN D. |
| KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : mündliche Offenbarung P : Zwischenliteratur T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus andern Gründen angeführtes Dokument * : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument | | | |

EPO FORM 150 (3.12.92) (P400)



Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung

EP 93 10 8104
PAGE2

| EINSCHLÄGIGE DOKUMENTE | | | |
|--|---|---|--|
| Kategorie | Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile | Betrifft Anspruch | KLASSIFIKATION DER ANMELDUNG (Int. Cl.5) |
| A | BIOSIS PREVIEWS DATABASE, Philadelphia, Ab.No.87107948 KLETZIN A 'COUPLED ENZYMATIC PRODUCTION...' * Zusammenfassung * & J BACTERIOL 171 (3). 1989. 1638-1643 ----- | 1,2 | |
| P,X | NUCLEIC ACIDS RESEARCH Bd. 20, Nr. 20, 1992, ARLINGTON, VIRGINIA US Seiten 5389 - 5396 Kletzin A. 'Molecular characterisation of a DNA ligase gene of the extremely thermophilic archaeon Desulfurolobus ambivalens shows close phylogenetic relationship to eukaryotic ligases' * das ganze Dokument * | 1-14 | |
| | | | RECHERCHIERTE SACHGEBIETE (Int. Cl.5) |
| Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt | | | |
| Recherchemat BERLIN | | Abschlußdatum der Recherche 11 AUGUST 1993 | Prüfer GURDJIAN D. |
| KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : mündliche Offenbarung P : Zwischenliteratur T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus andern Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument | | | |

EPO FORM 1503 03.82 (P0403)